ABDUL LATIF JAMEEL
Poverty Action Lab

TRADUCIENDO LA INVESTIGACIÓN EN ACCIÓN

El tamaño muestral en evaluaciones aleatorias

Guillermo Cruces

CEDLAS-Universidad Nacional de La Plata

Agenda del curso

1. Evaluaciones de impacto: ¿Qué son? ¿Por qué hacerlas? ¿Cuándo hacerlas?
2. Marco lógico, indicadores y medición del impacto
3. ¿Por qué aleatorizar?
4. ¿Cómo aleatorizar?
5. El tamaño muestral
6. Implementar una evaluación
7. Análisis e inferencia

Agenda del curso

1. Evaluaciones de impacto: ¿Qué son? ¿Por qué hacerlas? ¿Cuándo hacerlas?
2. Marco lógico, indicadores y medición del impacto
3. ¿Por qué aleatorizar?
4. ¿Cómo aleatorizar?
5. **El tamaño muestral**
6. Implementar una evaluación
7. Análisis e inferencia

RESUMEN Y OBJETIVOS

Resumen de la exposición

- Introducción al método científico
- Repaso: Estimación, prueba de hipótesis significancia estadística
- Poder:
  - Tamaño muestral
  - Magnitud del efecto
  - Otros factores que influyen en el poder: Varianza, aglomerados y otros
- Algunos aspectos prácticos

Objetivos de la clase

1. Refrescar/entender elementos básicos de estadística:
   - Estimación, Prueba de hipótesis, Significancia estadística
2. Concepto clave en evaluaciones aleatorias: Poder y los factores que lo influyen
   - Tamaño muestral, efectos mínimos detectables
   - Aglomerados y otros factores
3. Asignar recursos escasos en una evaluación.
   - Maximizar la probabilidad de encontrar efectos de programas que funcionan.

MÉTODO CIENTÍFICO

Método científico: Propuesta J-PAL

- Aplicación del método científico a las ciencias sociales

- El método científico implica:
  1. Proponer una hipótesis
  2. Diseñar estudios experimentales para probar la hipótesis

Aleatorización

- Los métodos experimentales solucionan el problema fundamental de la identificación: no podemos observar la misma unidad tratada y sin tratamiento simultáneamente.
- La aleatorización elimina el sesgo, pero no elimina el ruido.
- No basta con encontrar dos unidades equivalentes y comparar resultados…
- Funciona por la ley de los grandes números.

Planeando la aleatorización

- La identificación requiere la comparación de resultados esperados para dos *grupos, de manera creíble.*
- Usaremos la estadística para realizar las pruebas de hipótesis.
- Pero para esto necesitamos datos…
- Cuántos datos?

ESTIMACIÓN, PRUEBA DE HIPÓTESIS Y SIGNIFICANCIA ESTADÍSTICA

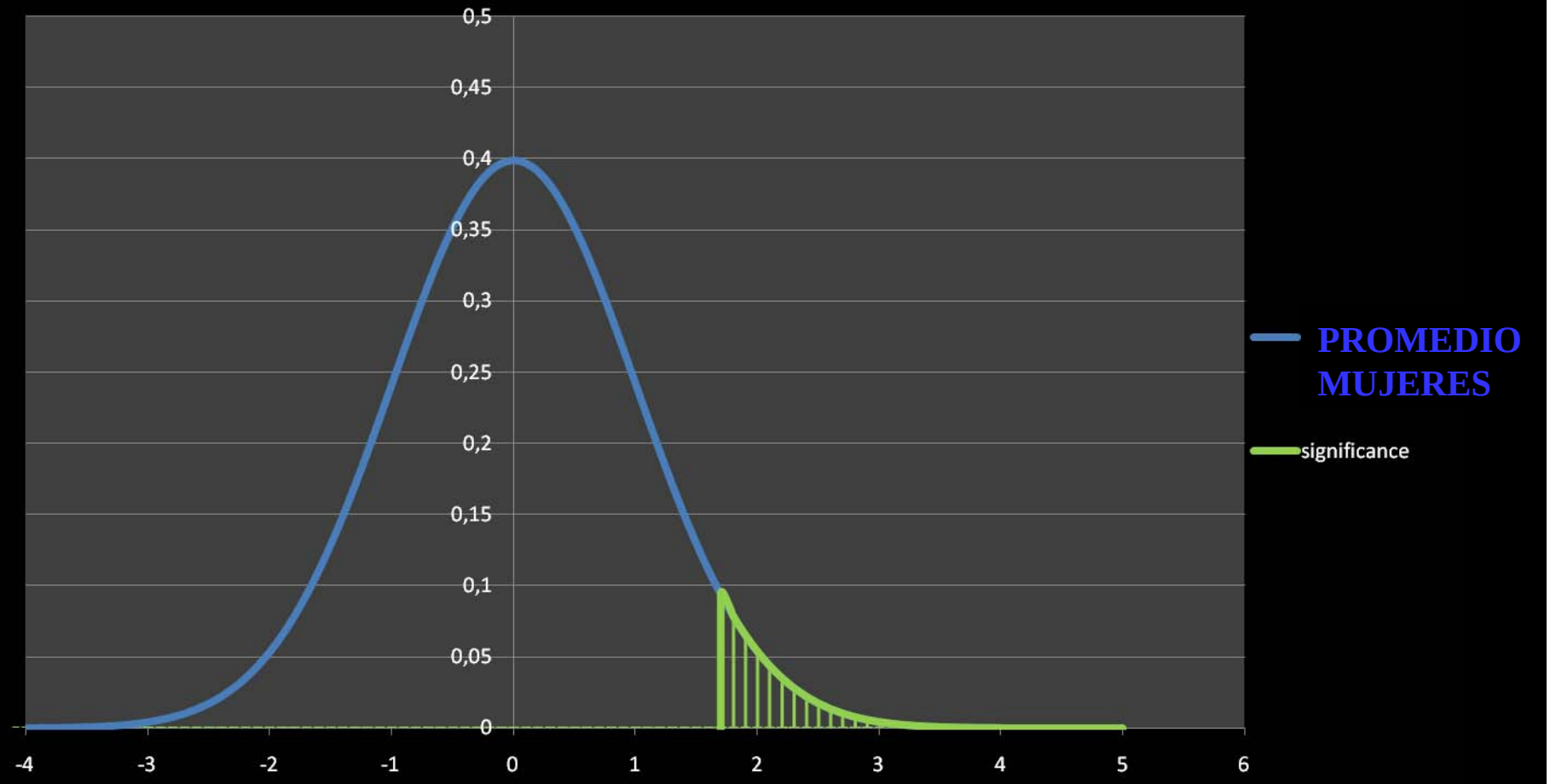
Prueba de hipótesis

- ¿Cómo probamos las hipótesis en ciencias sociales?
- Realizamos pruebas (*tests*) de hipótesis.
- En el caso de un tratamiento y dos grupos:
  - Diferencia en promedios de resultados
  - Otros momentos
  - Con paneles: diferencias en cambios

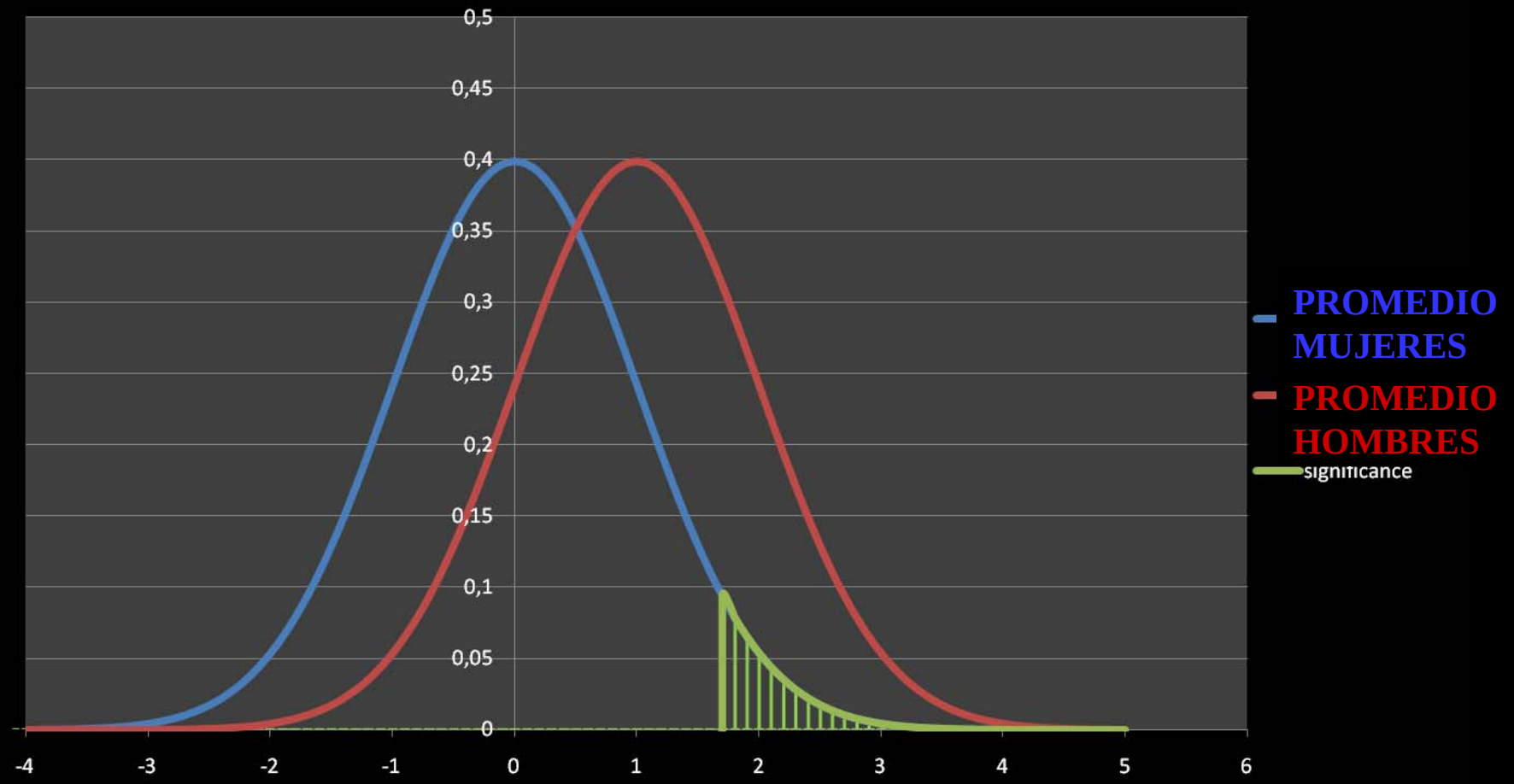
Prueba de hipótesis

- En estadística aplicada/evaluación, en lugar de la “presunción de inocencia” en la justicia, la norma es: “presunción de no diferencia”
- Hipótesis del evaluador/investigador:
  - No hay diferencia en estatura promedio entre hombres y mujeres
  - No hay diferencia en el resultado Y entre beneficiarios y no beneficiarios del programa X
- La evidencia debe demostrar lo contrario
- Hipótesis que testeamos en general: efecto=0

Ejemplo: ¿son en promedio más altos los hombres que las mujeres?



Ejemplo: ¿son en promedio más altos los hombres que las mujeres?



Ejemplo: Programa simulado.
Efecto en participación laboral - población

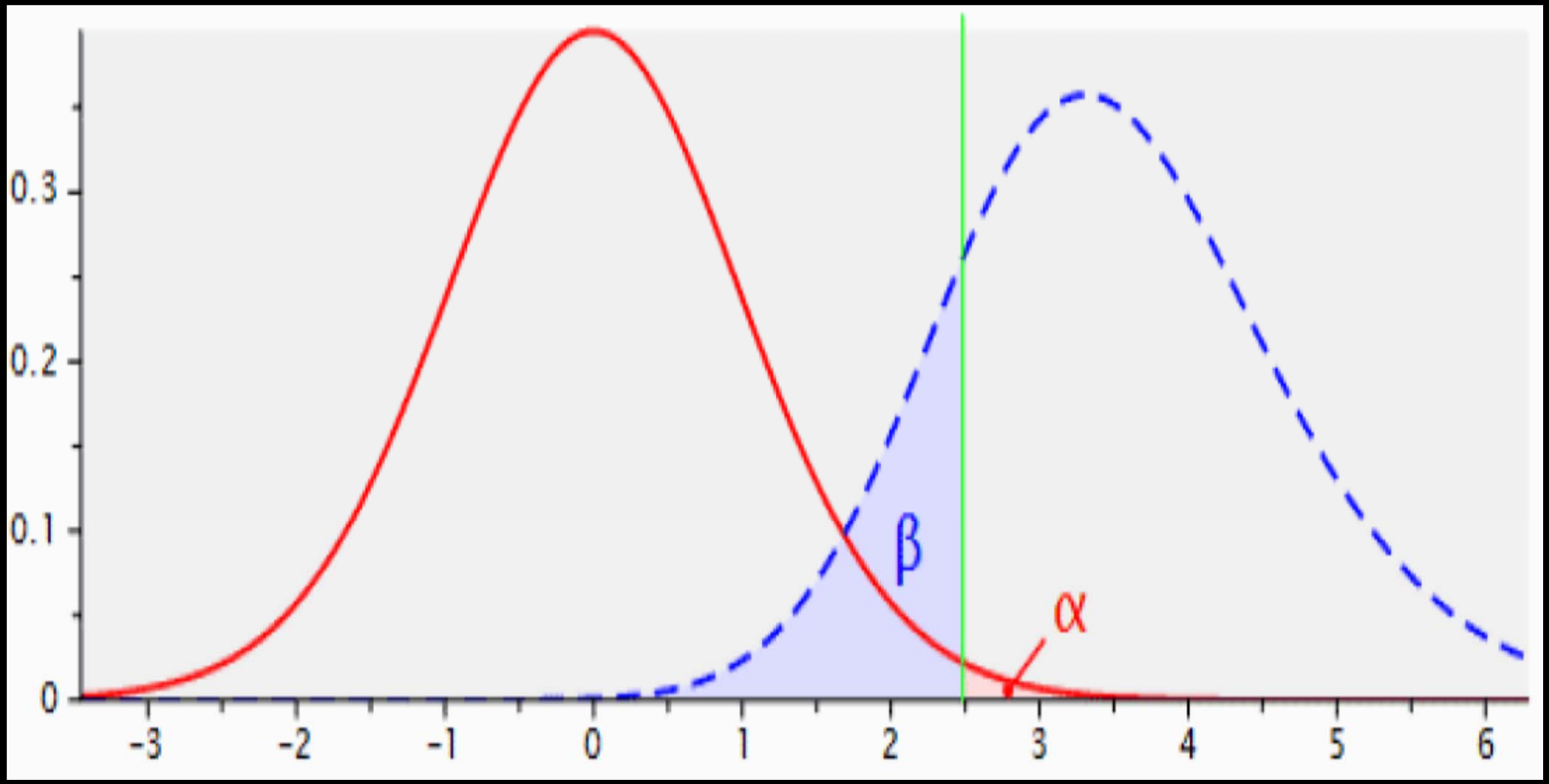
```
*Tenemos toda la población: 1.000.000
. tab y025 if d
       y025 |      Freq.     Percent        Cum.
------------+-----------------------------------
          0 |    237,500       47.50       47.50
          1 |    262,500       52.50      100.00
------------+-----------------------------------
. tab y025 if !d
       y025 |      Freq.     Percent        Cum.
------------+-----------------------------------
          0 |    250,000       50.00       50.00
          1 |    250,000       50.00      100.00
------------+-----------------------------------

*Efecto del tratamiento en la población
------------------------------------------------
        y025 |      Coef.   Std. Err.      t
-------------+----------------------------------
           d |        .025   .0009994    25.02
       _cons |          .5   .0007067   707.55
------------------------------------------------
```

Ejemplo: Programa simulado. Muestra de 10000

```
. Gen s=uniform()
. tab y025 if d & s<0.01
       y025 |      Freq.     Percent        Cum.
------------+-----------------------------------
          0 |      2,412       47.88       47.88
          1 |      2,626       52.12      100.00
------------+-----------------------------------
. tab y025 if !d & s<0.01
       y025 |      Freq.     Percent        Cum.
------------+-----------------------------------
          0 |      2,506       49.94       49.94
          1 |      2,512       50.06      100.00
------------+-----------------------------------
. reg y025 d if s<0.01
------------------------------------------------
        y025 |      Coef.   Std. Err.      t
-------------+----------------------------------
           d |   .0206407   .0099686     2.07
       _cons |   .5005978   .0070559    70.95
------------------------------------------------
Intervalo 95%: .0011003    .0401812
```

Estimación

- El efecto estimado es sólo válido para nuestra muestra. Cada muestra dará una respuesta diferente.
- Cómo usamos nuestra muestra para realizar afirmaciones sobre toda la población?
- Un intervalo de confianza del 95% nos dice que para el 95% de las muestras de ese tamaño que tomemos de esa población, el efecto estimado hubiera caído en ese intervalo.
- Está el cero en ese intervalo?
- El error estándar del estimador captura tanto el tamaño de la muestra como la variabilidad del resultado.
- Regla: IC de 95% es Efecto+/-2*Error.


0.3
0.2
0.1
0
-3
-2
-1
0
1
2
3
4
5
6
β
α

Estimación inexacta – variabilidad muestral, etc. Errores que podemos cometer

- En general, hipótesis nula=no hay efecto (en gral., queremos *rechazarla*).

- Tipo I: Rechazar una hipótesis nula verdadera (lobo!).
  - *5% implica 1 de 20*
- Tipo II: No rechazamos una hipótesis nula falsa.

- Para nuestras intervenciones: Tipo II implica que el programa tiene impacto pero no lo hallamos!

El problema es que podemos cometer error de dos tipos

		USTED CONCLUYE	
		Rechazo H0 - Programa tiene efecto	*No rechazo H0 - Programa no tiene efecto*
LA VERDAD	*Programa tiene efecto*	☺	Error tipo II (bajo poder) ☹
	Programa no tiene efecto	Error tipo I ☹	☺

El problema es que podemos cometer error de dos tipos

		USTED CONCLUYE	
		Rechazo H0 - Programa tiene efecto	*No rechazo H0 - Programa no tiene efecto*
LA VERDAD	*Programa tiene efecto*	☺	Error tipo II (bajo poder) ☹
	Programa no tiene efecto	Error tipo I ☹	☺

Probabilidad de rechazar la hipótesis nula dado que es verdadera

El problema es que podemos cometer error de dos tipos

		USTED CONCLUYE	
		Rechazo H0 - Programa tiene efecto	*No rechazo H0 - Programa no tiene efecto*
LA VERDAD	*Programa tiene efecto*	☺	Error tipo II (bajo poder) ☹
	Programa no tiene efecto	Error tipo I ☹	☺

Probabilidad de NO rechazar la hipótesis nula dado que es falsa

PODER:

Tamaño muestral

La tiranía del poder – Bland, BMJ 2008

- La idea del poder estadístico es decepcionantemente sencilla.
- Vamos a realizar un estudio donde evaluaremos la evidencia usando pruebas de hipótesis.
- Decidimos entonces cuán grande es el efecto que queremos detectar, es decir, qué tamaño de efecto valdría la pena conocer.
- A partir de eso elegimos un tamaño muestral de manera que, si este fuera el efecto real en la población, una alta proporción de las muestras posibles producirían una diferencia estadísticamente significativa.

La pregunta de hoy

- Poder: probabilidad de que una prueba de hipótesis rechazará la hipótesis nula cuando ésta es *falsa* – probabilidad de no cometer un error de tipo II.
- Un poder del 80% nos dice que en 80% de los experimentos con este tamaño de muestra realizados en esta población, si hay un efecto, lo encontraremos en nuestra muestra para el nivel de significatividad que definimos.

La pregunta de hoy

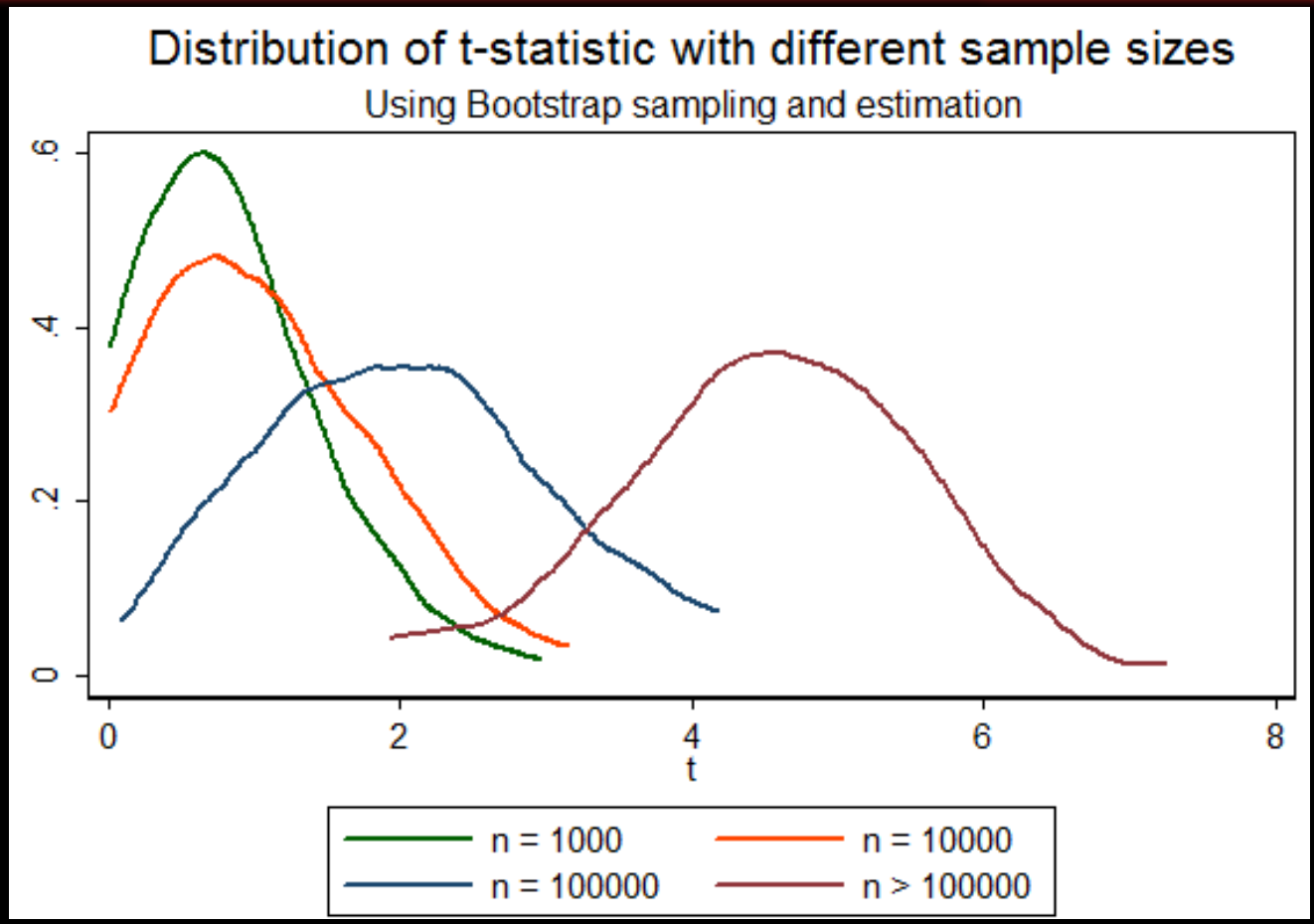
- Entre otras cosas, implica:

¿Qué tan grande debe ser la muestra para detectar “de manera creíble” una diferencia entre dos grupos?

Intuición

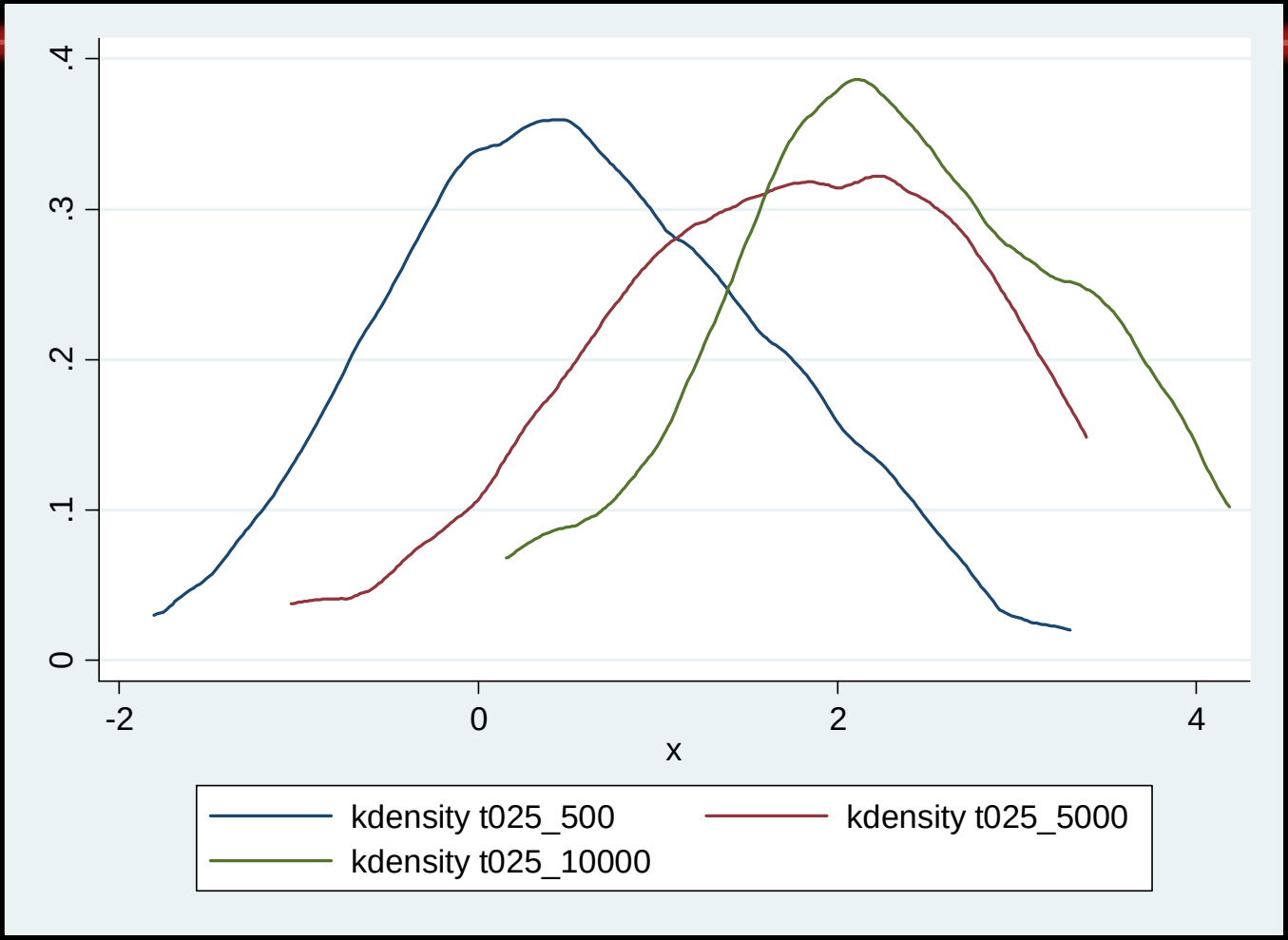
- Entre más grandes son nuestros grupos muestrales:
  - Más nos aproximamos a las características de la población, reducimos nuestra incertidumbre, y por tanto
  - Más probable concluir que hay una diferencia, dado que en la población si existe tal diferencia

Ejemplo de estimación con datos observacionales



Efecto de tener dos hijos del mismo sexo en la oferta laboral femenina. Censo de USA, 1980. Cruces, Tortarolo & Pinto, 2011

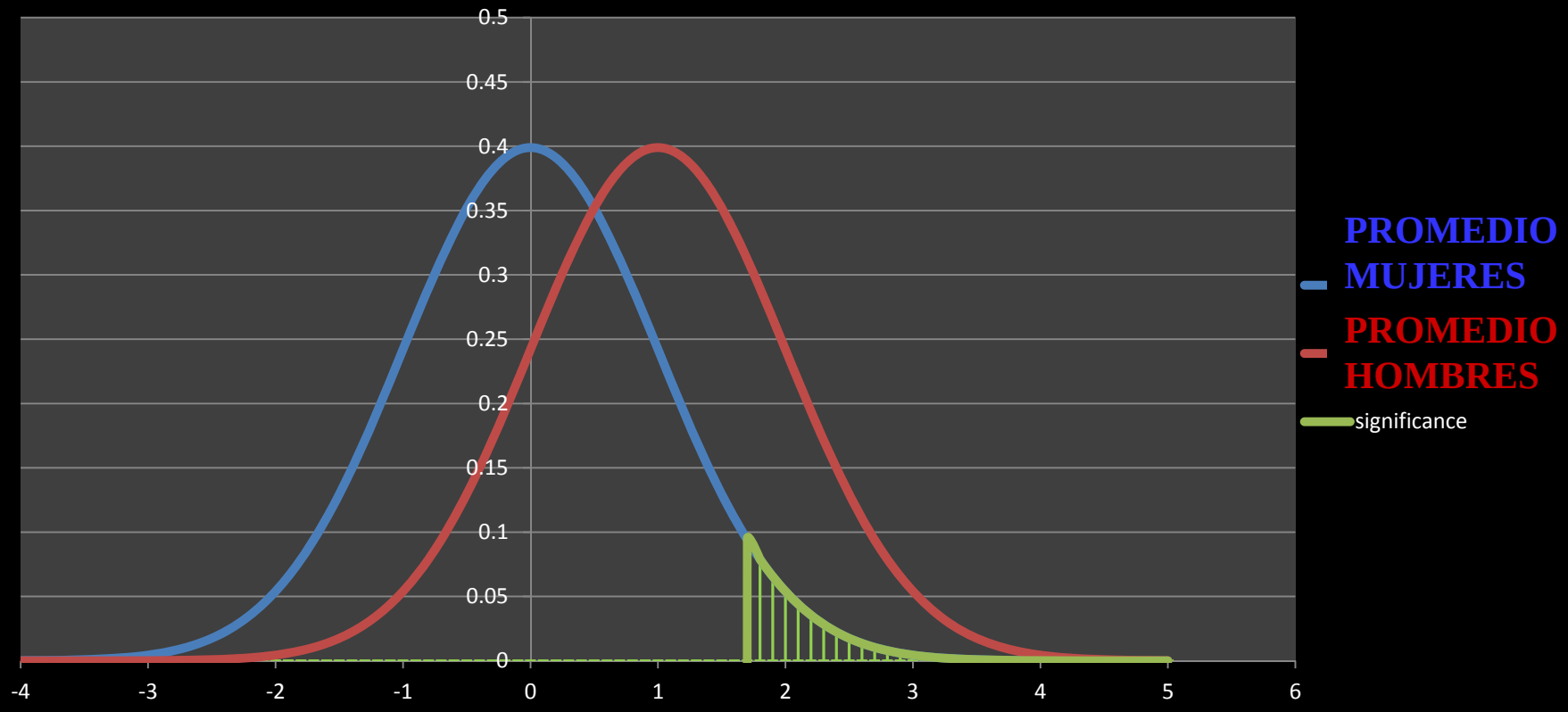
Simulaciones: efecto pequeño, muestras cada vez más grandes – estadístico *t* (>2?)



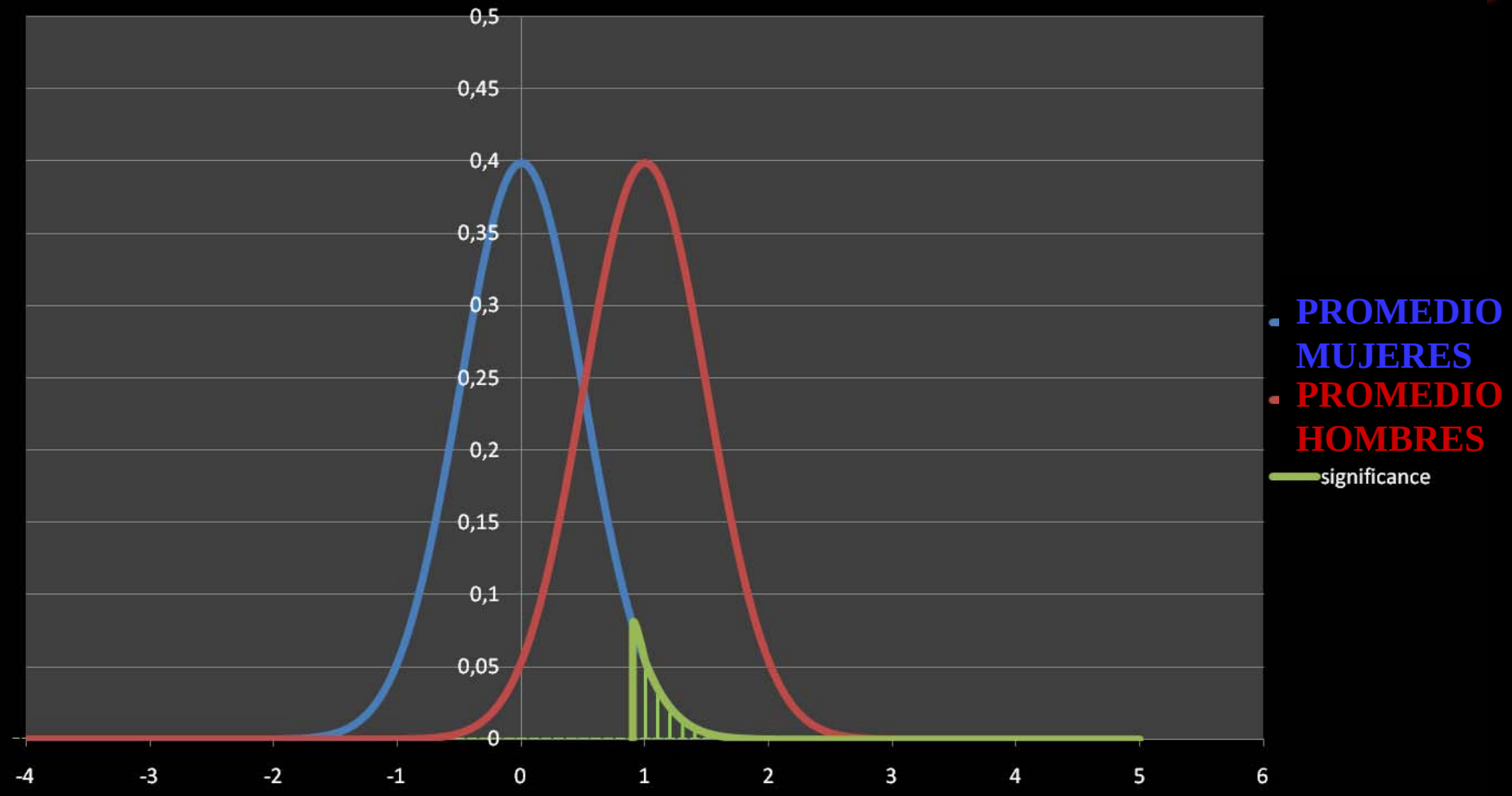
Mayor muestra, mayor significatividad a mismo efecto

¿Cómo cambia el poder con el tamaño muestral?

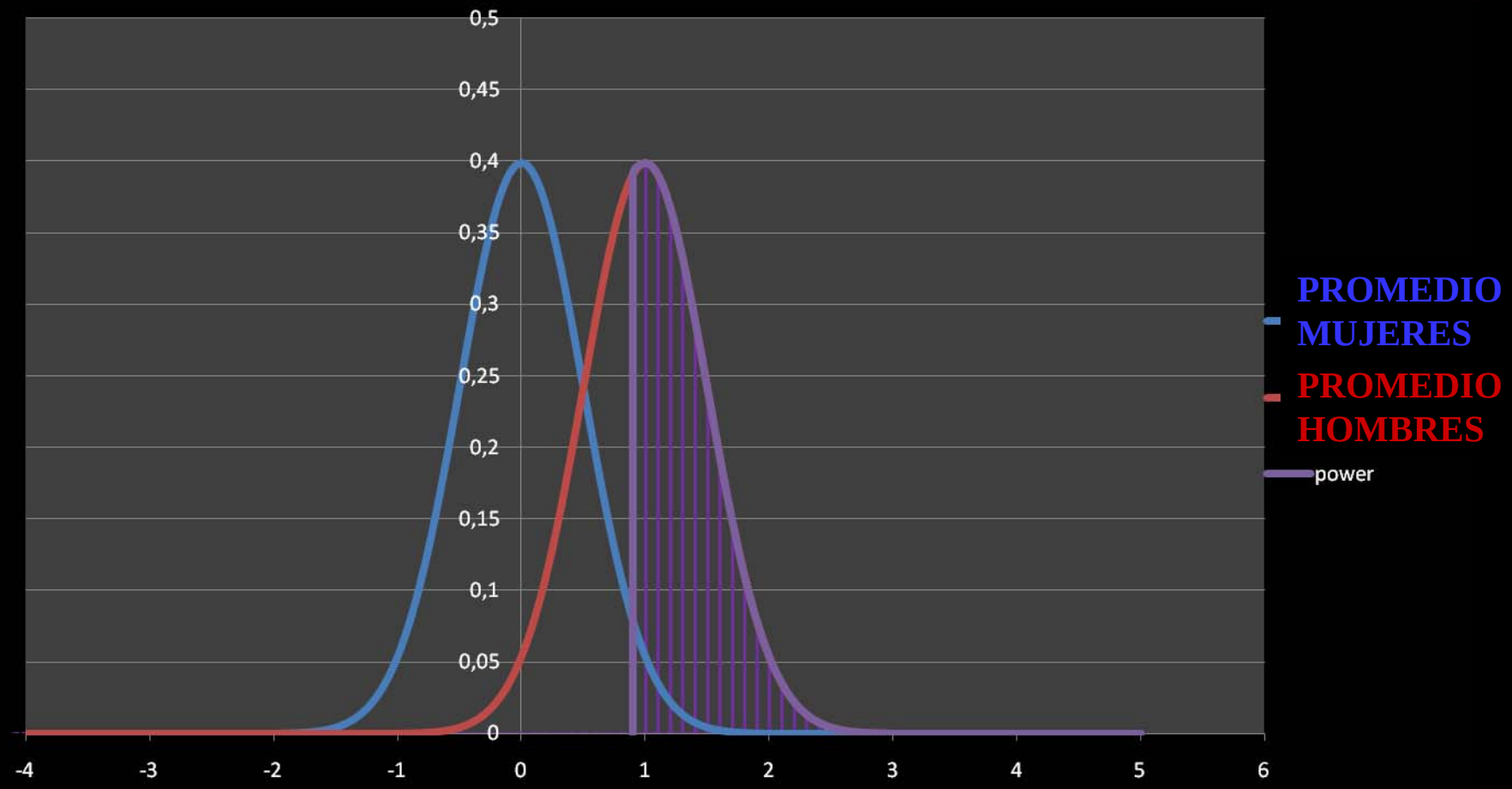
Dado un nivel de significancia y una magnitud de efecto…

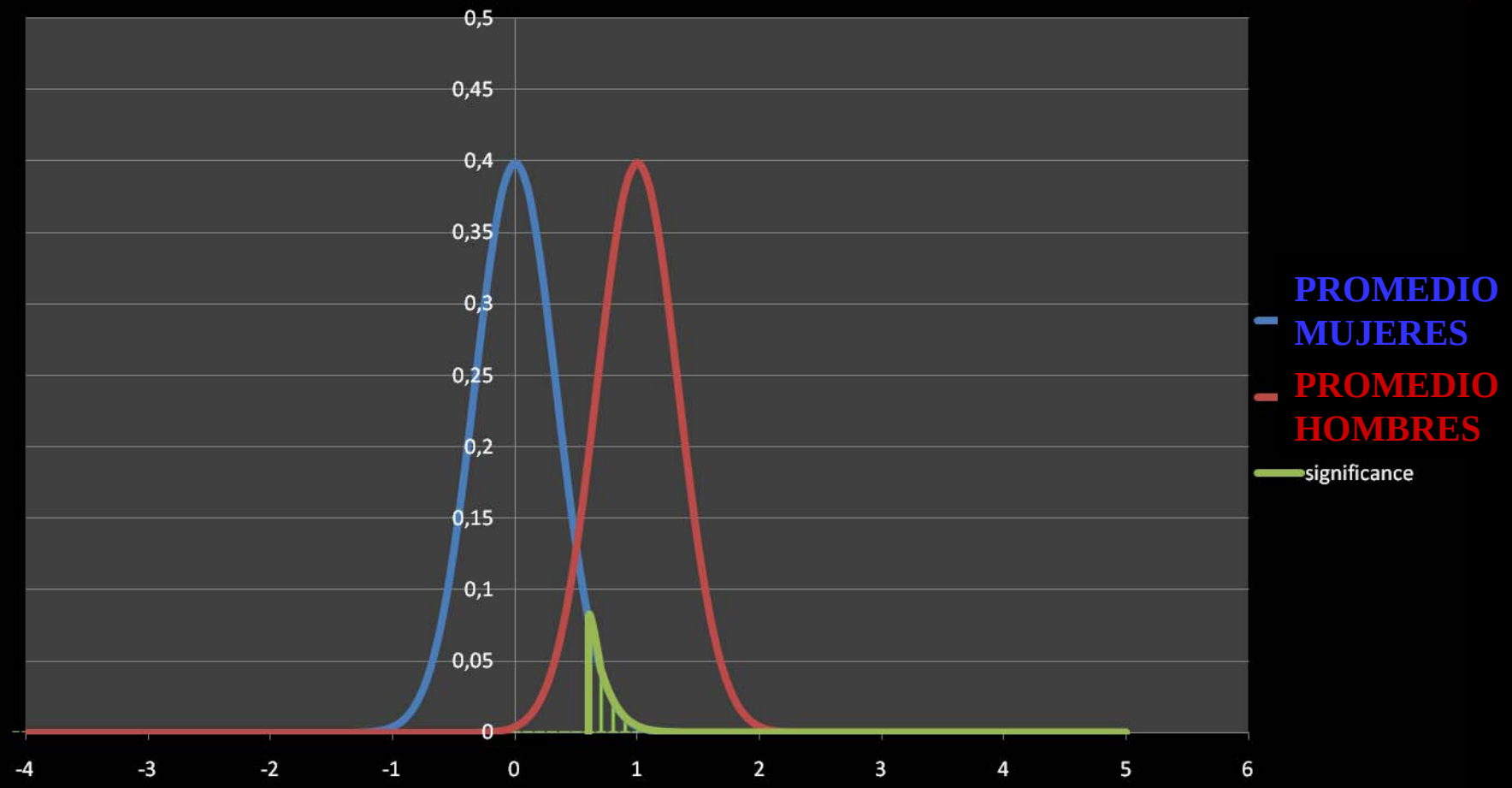


A mayor muestra, menor varianza de mi estimador, y por tanto mayor poder…

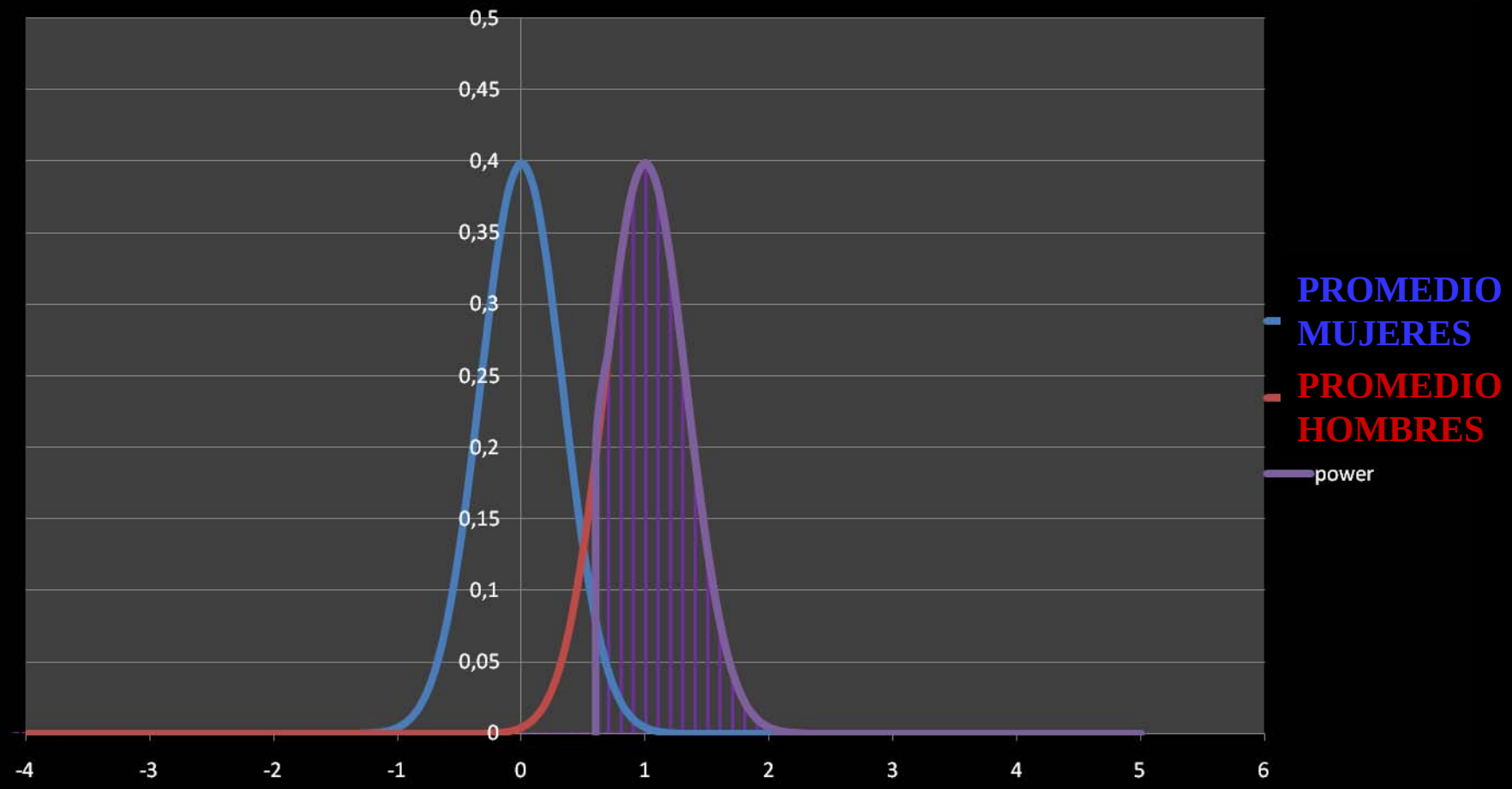


Poder: 64%


0,5
0,45
0,4
0,35
0,3
0,25
0,2
0,15
0,1
0,05
0
-4
-3
-2
-1
0
1
2
3
4
5
6
PROMEDIO MUJERES
PROMEDIO HOMBRES
power


0,5
0,45
0,4
0,35
0,3
0,25
0,2
0,15
0,1
0,05
0
-4
-3
-2
-1
0
1
2
3
4
5
6
PROMEDIO MUJERES
PROMEDIO HOMBRES
significance

Poder: 91%


0,5
0,45
0,4
0,35
0,3
0,25
0,2
0,15
0,1
0,05
0
-4
-3
-2
-1
0
1
2
3
4
5
6
PROMEDIO MUJERES
PROMEDIO HOMBRES
power

En resumen, hasta ahora

- Dos tipos de errores de decisión:
- Siempre tenemos control sobre el error de tipo I
  - Es el umbral de decisión que nosotros escojamos usualmente 5% (una/dos colas)
- NO siempre controlamos el error de tipo II,
  - Con encuentas ya realizadas, el número de observaciones está determinado

(aunque podemos calcular el Efecto Mínimo Detectable con esa muestra y los DE)

PODER

Magnitud del efecto y su variabilidad

En resumen, hasta ahora

- En diseños experimentales podemos determinar cuánto error de tipo II vamos a tolerar => Cálculos de poder

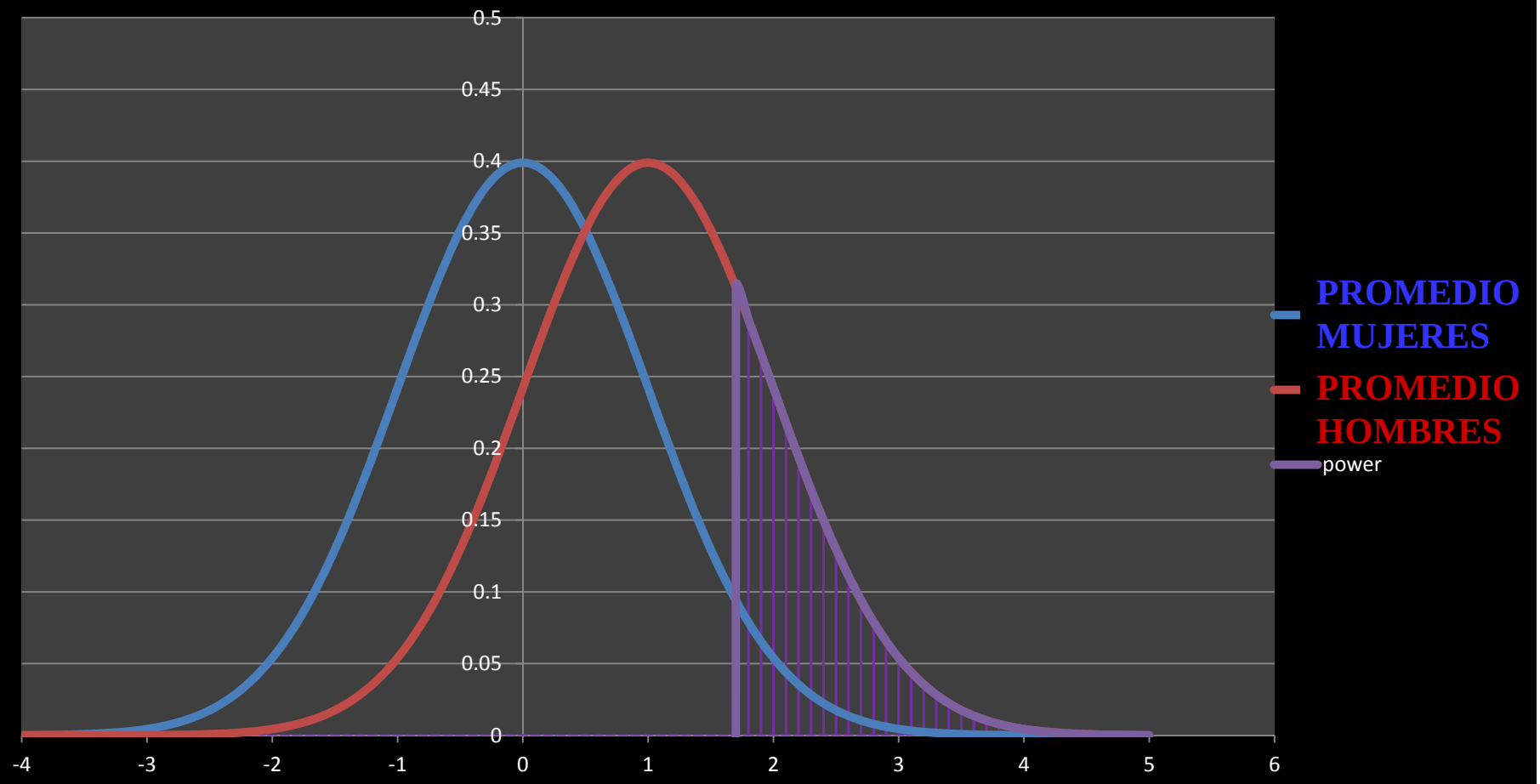
- Entre mayor sea la diferencia promedio entre los dos grupos, mayor el poder (menor el error de tipo II) dada la muestra que tenemos

¿Cómo elijo con anterioridad una magnitud de efecto?

- ¿Cuál es el efecto menor que justificaría el programa que se está realizando?
- Si el efecto es menor que eso, no nos interesaría mucho probar que sea diferente a cero – no es económicamente significativo…
- En contraste, si cualquier efecto más grande que ese justificara adoptar este programa: queremos poder distinguirlo de cero

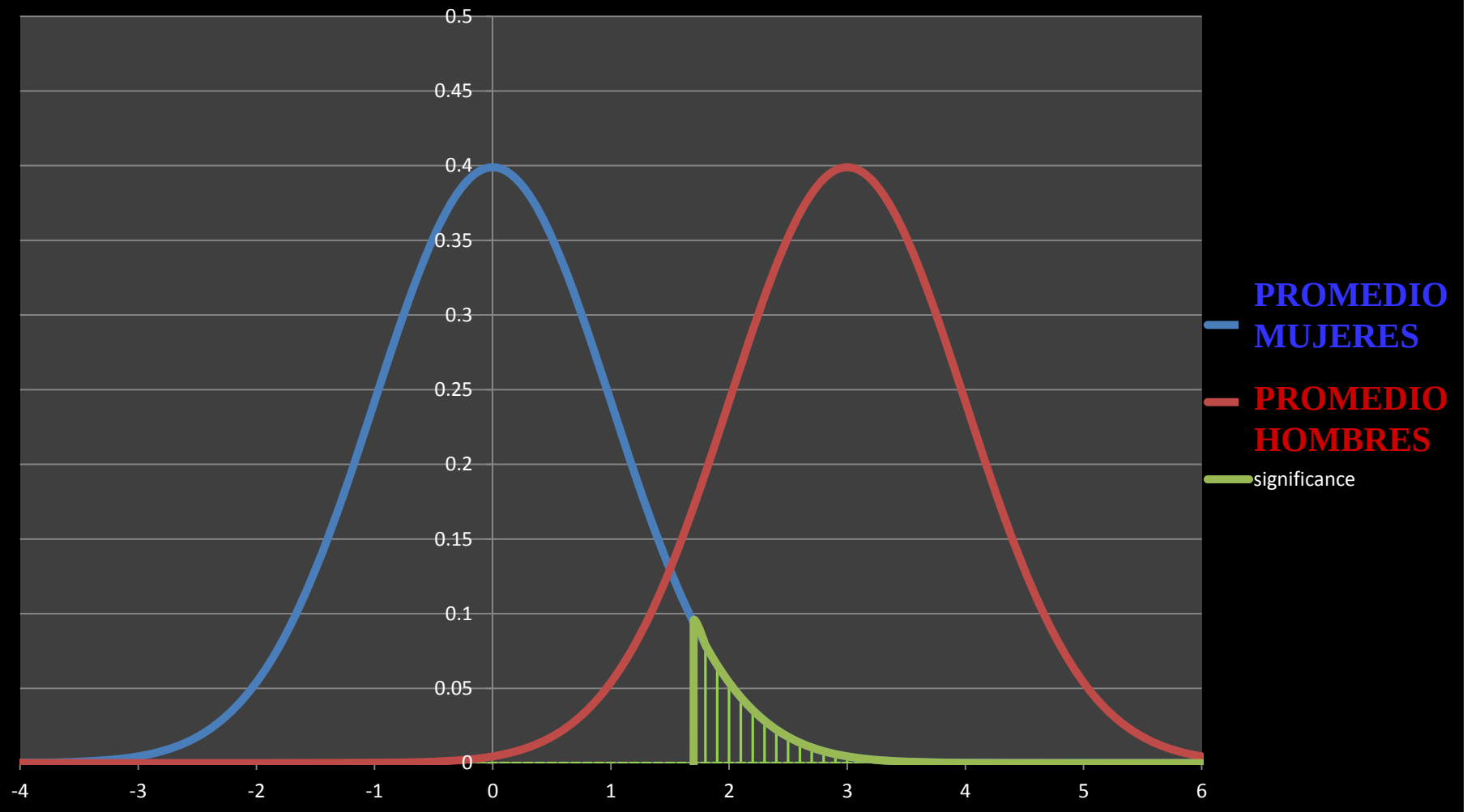
¿Magnitud de efecto en programas similares?

Si diferencia observada en estatura fuera de 1 DS…

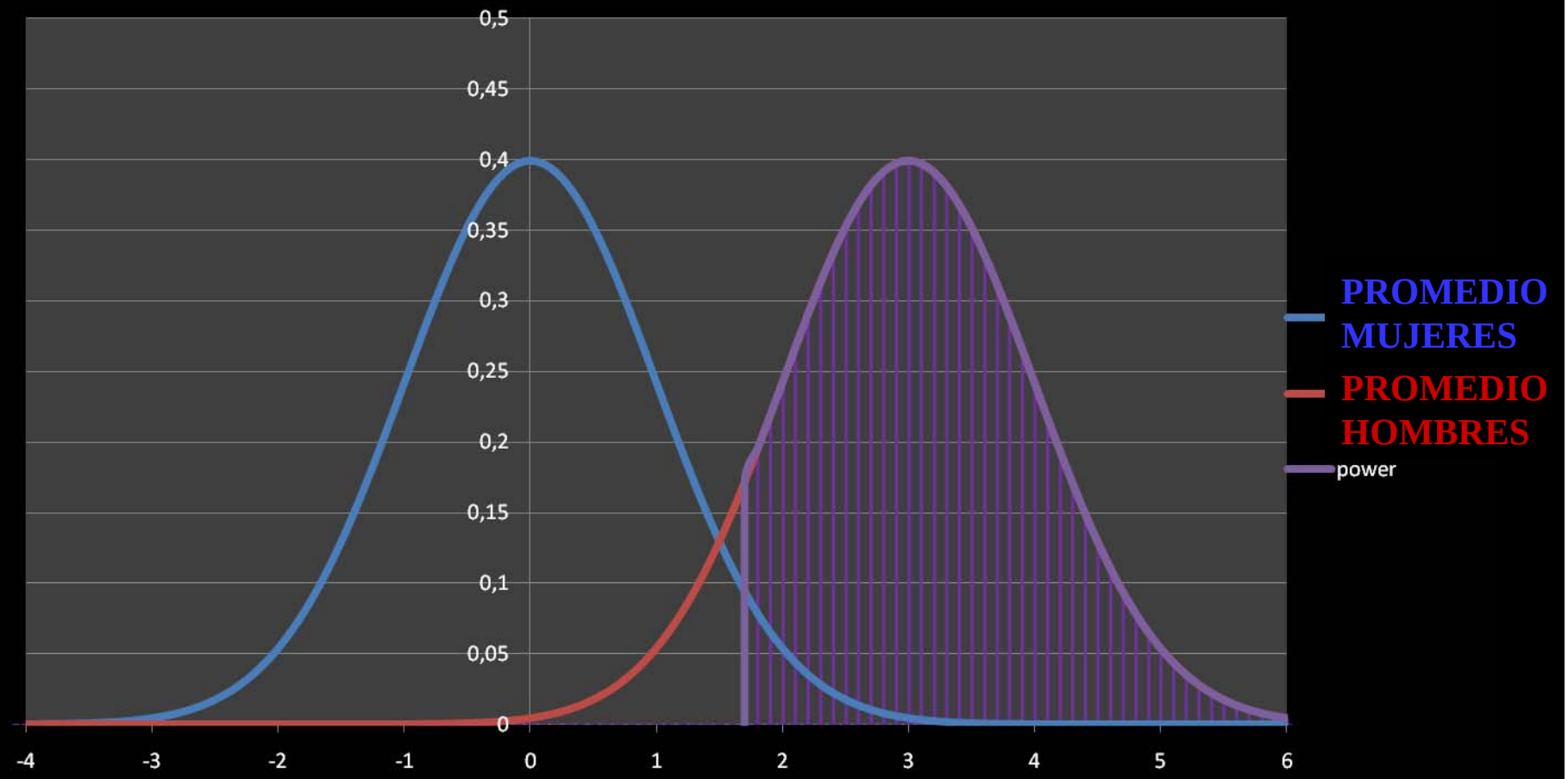


La hipótesis nula se rechazaría sólo el 26% de las veces

¿Si diferencia observada en estatura fuera 3 DS cuanto sería el poder?

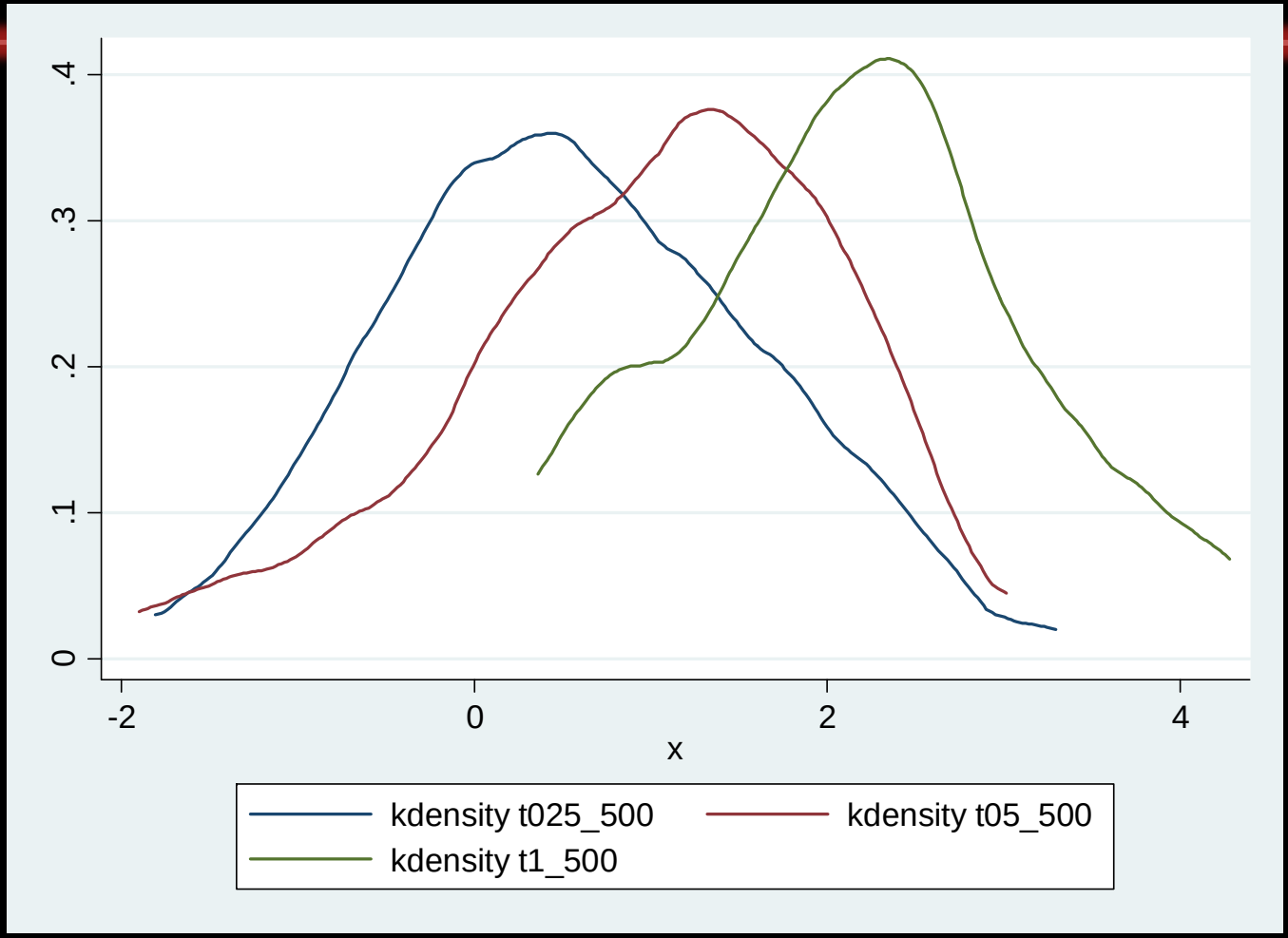


Poder: 91%



La hipótesis nula se rechazaría el 91% de las veces

Simulaciones: muestra fija, efectos cada vez más grandes – estadístico *t* (>2?)

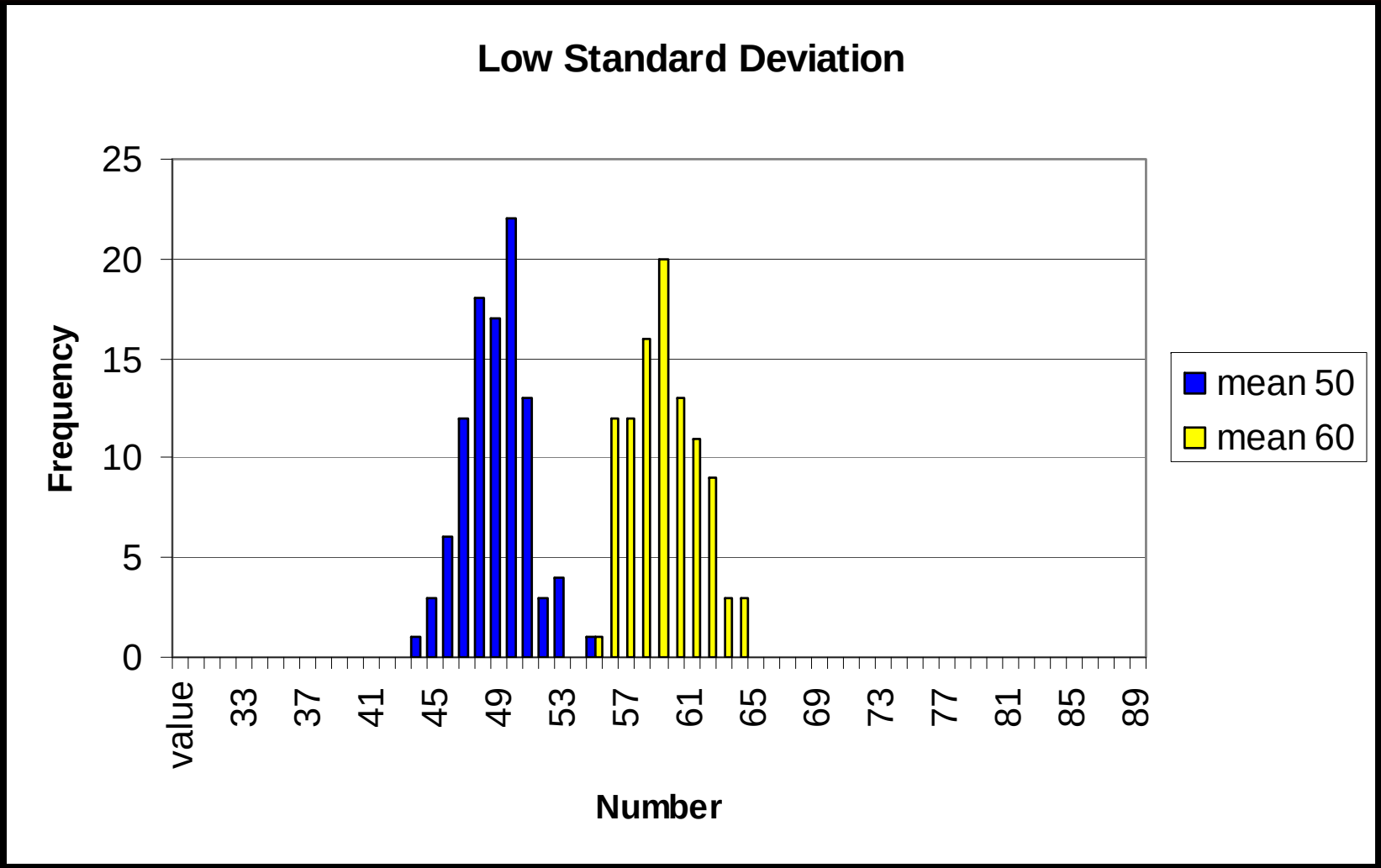


Mayor efecto, mayor significatividad a mismo N

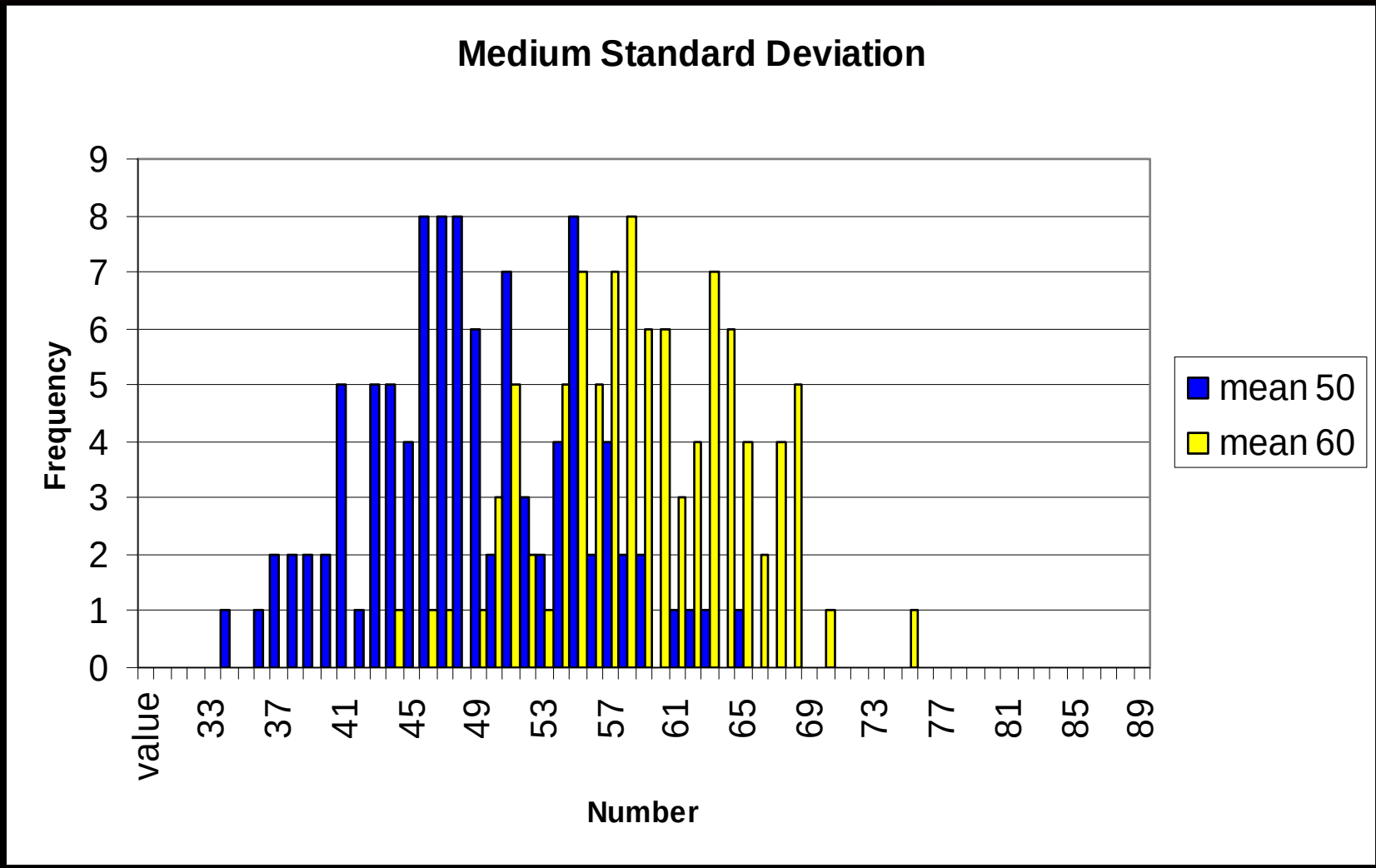
Efectos estandarizados

- Unidades típicas de medida son absolutas
  - Cms, puntos, ocurrencias etc.
- No es lo mismo una diferencia de 20cms cuando desviación estándar (variabilidad) es 20cms que cuando es 40cms
- Efecto estandarizado es magnitud del efecto dividida por desviación estándar de la variable de resultado
  - Sus unidades son desviaciones estándar

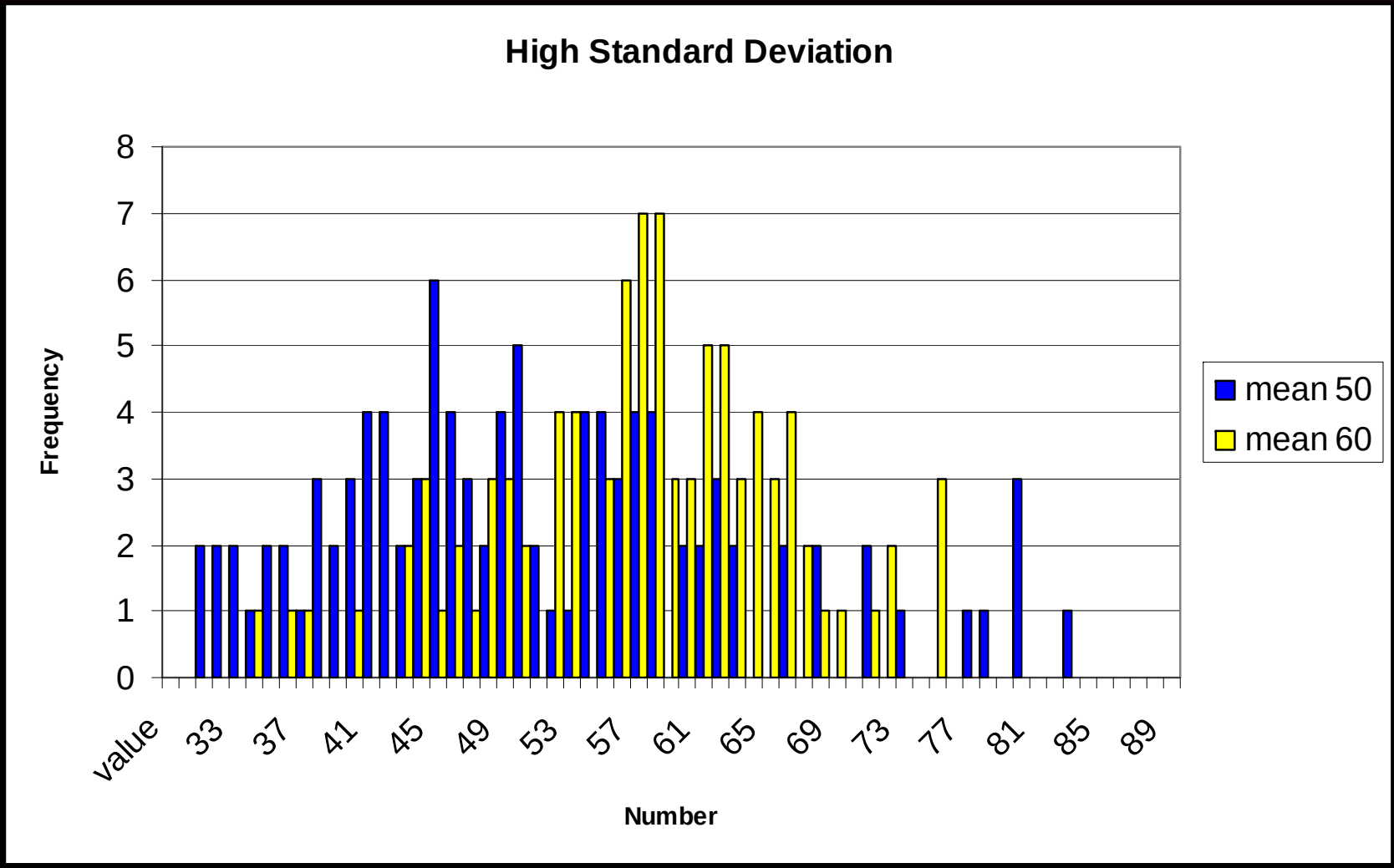
La importancia de la variabilidad de los resultados
Alta precisión, efectos claros:



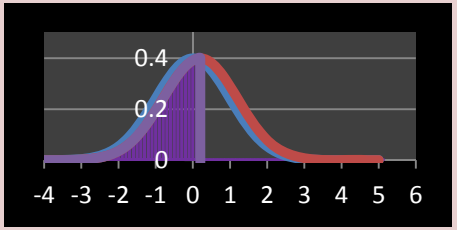
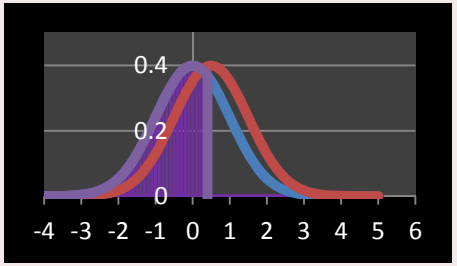
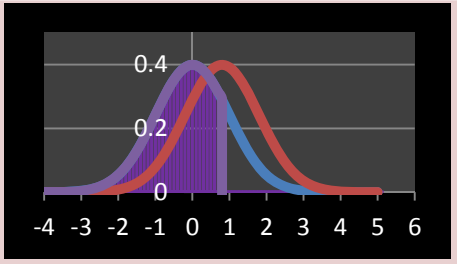
Menos precisión



Menos claro…


High Standard Deviation
Frequency
0
1
2
3
4
5
6
7
8
value
33
37
41
45
49
53
57
61
65
69
73
77
81
85
89
Number
mean 50
mean 60

Magnitud del efecto estandarizado

Una magnitud de efecto de…	Se considera…	…y significa que…	
0.2	Pequeña-modesta	El miembro promedio del grupo de tratamiento tuvo un mejor resultado que el **percentil 58** del grupo de control	
0.5	Modesta-grande	El miembro promedio del grupo de tratamiento tuvo un mejor resultado que el **percentil 69** del grupo de control	
0.8	Grande	El miembro promedio del grupo de tratamiento tuvo un mejor resultado que el **percentil 79** del grupo de control	

La pregunta de hoy

¿Qué tan grande debe ser la muestra para detectar “de manera creíble” una diferencia entre dos grupos?

PODER

Cálculo analítico

3 Ingredientes esenciales, hasta ahora...

Tamaño muestral

- Nivel de significancia
  5%, 10%
- Poder mínimo deseado
  80%, 90%
- Efecto mínimo estandarizado
  0.2 DS, 0.5 DS

Ingredientes para un cálculo de poder estadístico

Lo que necesitamos	**Dónde obtenerlo**
Nivel de significatividad	Convencional. 5%-1% etc.En cuanto más bajo, más grande la muestra para el mismo poder.
El nivel y la variabilidad del efecto	De encuestas anteriores en situaciones similares. A mayor variabilidad, mayor muestra necesaria para el mismo poder.
El efecto que queremos encontrar	Cuál es el menor efecto que justificaría una respuesta de política? En cuanto menor el efecto, mayor la muestra para el mismo poedr. Cuidado: Pensar en términos de la realidad – el efecto puede ser mayor a cero estadísticamente, pero irrelevante…

Una relación muy sencilla entre ellos…

$$N = 4(t_{1-\kappa} + t_{\alpha})^2 \cdot \frac{1}{MEE^2}$$

t_{1-k} = Valor crítico de t asociado a poder 1-k.
Para poder k=80%, t_{1-k}=0.84

t_{α} = Valor crítico de t asociado a nivel de significancia α.
Para α=0.05 t_{α}=1.666

MEE = Magnitud de Efecto Estandarizado
= Magnitud de Efecto / Desviación Estándar

¿Cómo cambia N cuando…?

$$N = 4(t_{1-\kappa} + t_{\alpha})^2 \cdot \left[\frac{DE}{ME}\right]^2$$

t_{1-k} = Valor crítico de t asociado a poder 1-k.
Para poder k=80%, t_{1-k}=0.84

t_{α} = Valor crítico de t asociado a nivel de significancia α.
Para α=0.05 t_{α}=1.666
= Magnitud de Efecto / Desviación Estándar

Ejercicio

- Imagine que usted quiere diseñar un experimento para estudiar el impacto de la tecnología (por ejemplo semillas de alto rendimiento) sobre la productividad agrícola
- Estime cuántos agricultores necesita si quiere encontrar un efecto en el rendimiento por hectárea de 10% dado un rendimiento promedio de 50 kilos/hectárea y una desviación estándar de 60 kilos/hectárea

¿Parámetros?

Ejercicio

- DE=60
- ME=10% de 50=5
- MEE =5/0.083

Ejercicio: Poder=0.8

```
. sampsi 50 55,sd(60) power(0.8) onesided

Estimated sample size for two-sample comparison of means

Test Ho: m1 = m2, where m1 is the mean in population 1
                    and m2 is the mean in population 2
Assumptions:

         alpha =   0.0500  (one-sided)
         power =   0.8000
            m1 =       50
            m2 =       55
           sd1 =       60
           sd2 =       60
         n2/n1 =     1.00

Estimated required sample sizes:

            n1 =     1781
            n2 =     1781
```

Ejercicio: Poder=0.8 - MEE

```
. sampsi 0 0.083333,power(0.8) onesided sd(1)

Estimated sample size for two-sample comparison of means

Test Ho: m1 = m2, where m1 is the mean in population 1
                    and m2 is the mean in population 2
Assumptions:

         alpha =   0.0500  (one-sided)
         power =   0.8000
            m1 =         0
            m2 =  .083333
           sd1 =         1
           sd2 =         1
         n2/n1 =      1.00

Estimated required sample sizes:

            n1 =      1781
            n2 =      1781
```

Si tuviéramos menos varianza…

```
. sampsi 50 55,sd(31.78) power(0.8) onesided

Estimated sample size for two-sample comparison of means

Test Ho: m1 = m2, where m1 is the mean in population 1
                    and m2 is the mean in population 2
Assumptions:

         alpha =   0.0500  (one-sided)
         power =   0.8000
            m1 =       50
            m2 =       55
           sd1 =    31.78
           sd2 =    31.78
         n2/n1 =     1.00

Estimated required sample sizes:

            n1 =      500
            n2 =      500
```

Si el efecto fuera el triple…

```
. sampsi 50 65,sd(60) power(0.8) onesided

Estimated sample size for two-sample comparison of means

Test Ho: m1 = m2, where m1 is the mean in population 1
                    and m2 is the mean in population 2
Assumptions:

         alpha =   0.0500  (one-sided)
         power =   0.8000
            m1 =       50
            m2 =       65
           sd1 =       60
           sd2 =       60
         n2/n1 =     1.00

Estimated required sample sizes:

            n1 =      198
            n2 =      198
```

Si aceptamos menos poder: 0.7

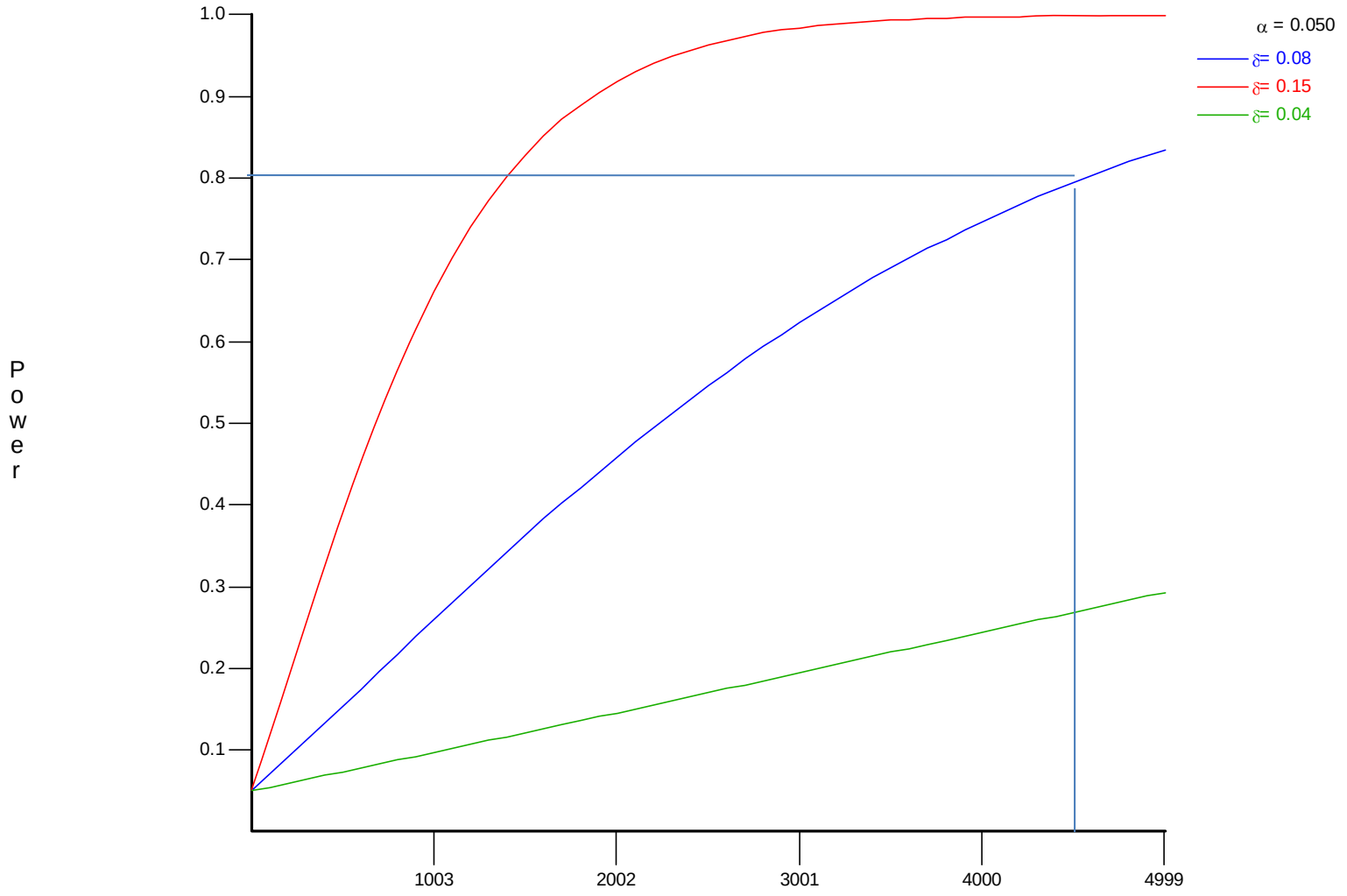
```
. sampsi 50 55,sd(60) power(0.7) onesided

Estimated sample size for two-sample comparison of means

Test Ho: m1 = m2, where m1 is the mean in population 1
                    and m2 is the mean in population 2
Assumptions:

         alpha =   0.0500  (one-sided)
         power =   0.7000
            m1 =       50
            m2 =       55
           sd1 =       60
           sd2 =       60
         n2/n1 =     1.00

Estimated required sample sizes:

            n1 =     1356
            n2 =     1356
```

¿Y si ahorramos en la encuesta?
Con 1000 casos seguro estamos bien… ¿o no?

```
. sampsi 50 55,sd(60) n(500) a(0.05) onesided

Estimated power for two-sample comparison of means

Test Ho: m1 = m2, where m1 is the mean in population 1
                    and m2 is the mean in population 2
Assumptions:

         alpha =   0.0500  (one-sided)
            m1 =       50
            m2 =       55
           sd1 =       60
           sd2 =       60
sample size n1 =      500
            n2 =      500
         n2/n1 =     1.00

Estimated power:

         power =   0.3717
```

Poder y tamaño de muestra

Programa: Optimal Design 3.0. NB: Diferencia: two sided, N1=N2=2261



PODER

Otros aspectos

Tres consideraciones prácticas que afectan los requerimientos muestrales

1. ¿Hay datos de línea de base?

2. ¿La asignación al tratamiento es a nivel individual o a nivel grupal (por ejemplo colegio o municipio)?

3. ¿Se cumple total o parcialmente con el protocolo de asignación experimental?

1. ¿Hay datos de línea de base?

- Dada una magnitud de efecto esperado y un tamaño muestral, hay mayor poder si hay covariables de línea de base que tengan poder explicativo sobre indicador de resultado
- Dicho de otra forma, dado una MEE y poder deseado, requiero menos tamaño muestral para detectarlo (menos $$$)
- Lo que importa es la varianza residual – después de eliminar el efecto de la covariable relevante.

1. ¿Hay datos de línea de base?

- La co-variable de línea de base con mayor poder explicativo es el indicador de impacto medido en línea de base

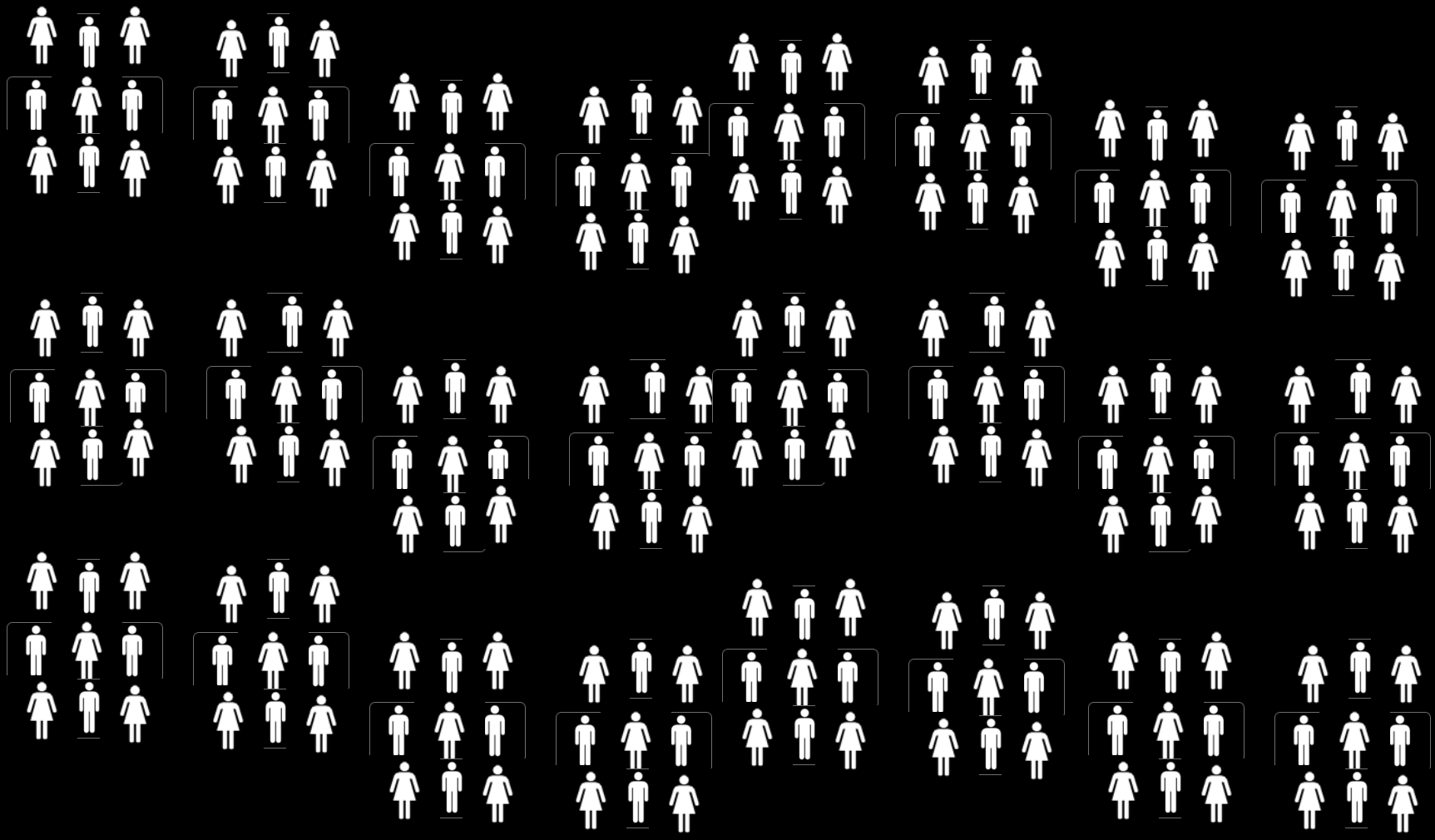
- Por ejemplo, si quiero medir impacto de programa educativo sobre puntaje en pruebas, es buena idea administrar el examen también en línea de base

2. ¿AsignaciÓn a nivel individual o a nivel de grupo? Conglomerados

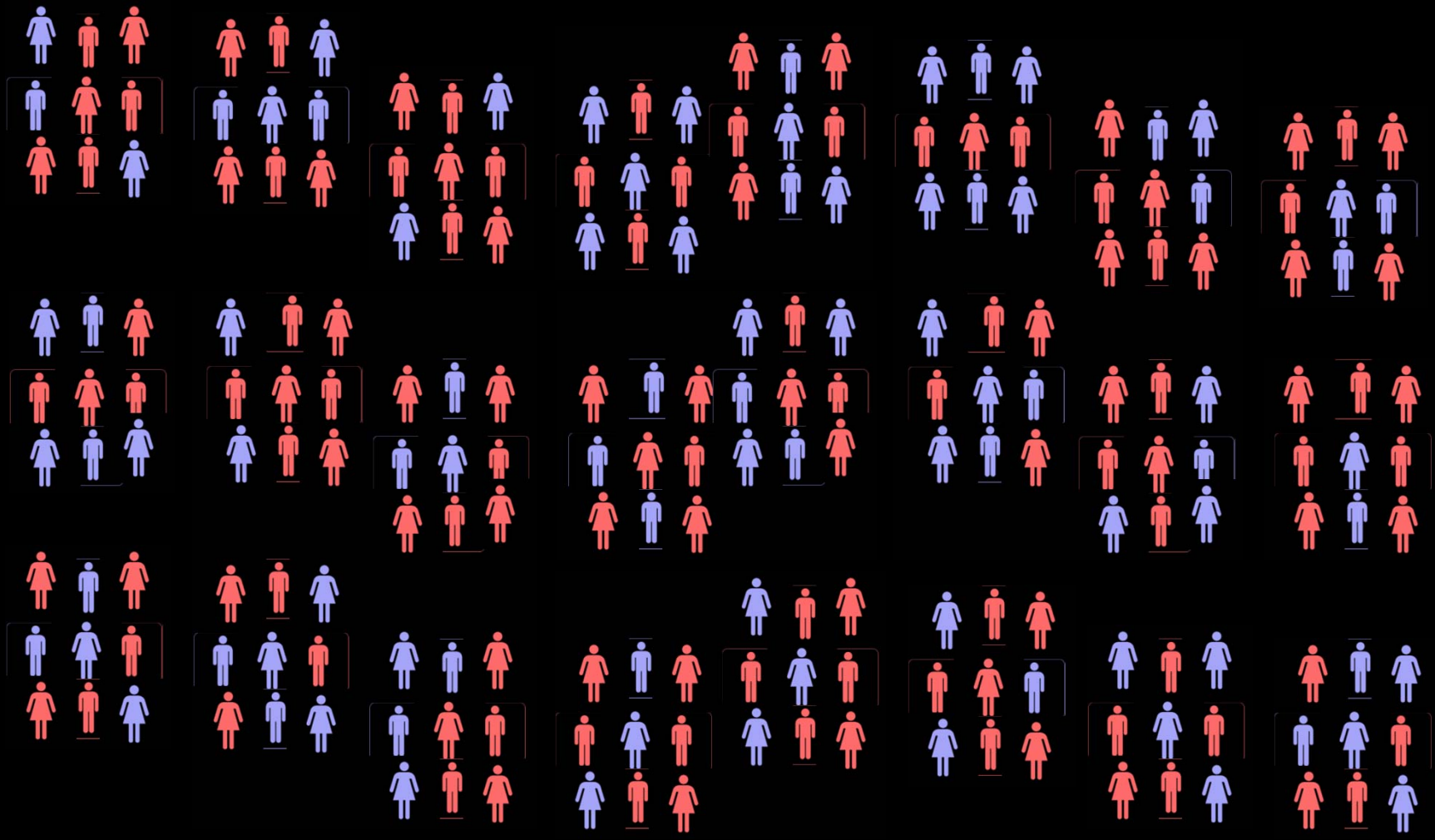
- Diseños en conglomerados son experimentos en que unidades sociales o conglomerados en vez de personas se asignan aleatoriamente a grupos de intervención

- La unidad de aleatorización (por ejemplo, el colegio) es más amplia que la unidad de análisis (por ejemplo, estudiantes)
  - Aleatorizar a nivel de colegio y utilizar pruebas a nivel de niño como indicador de impacto

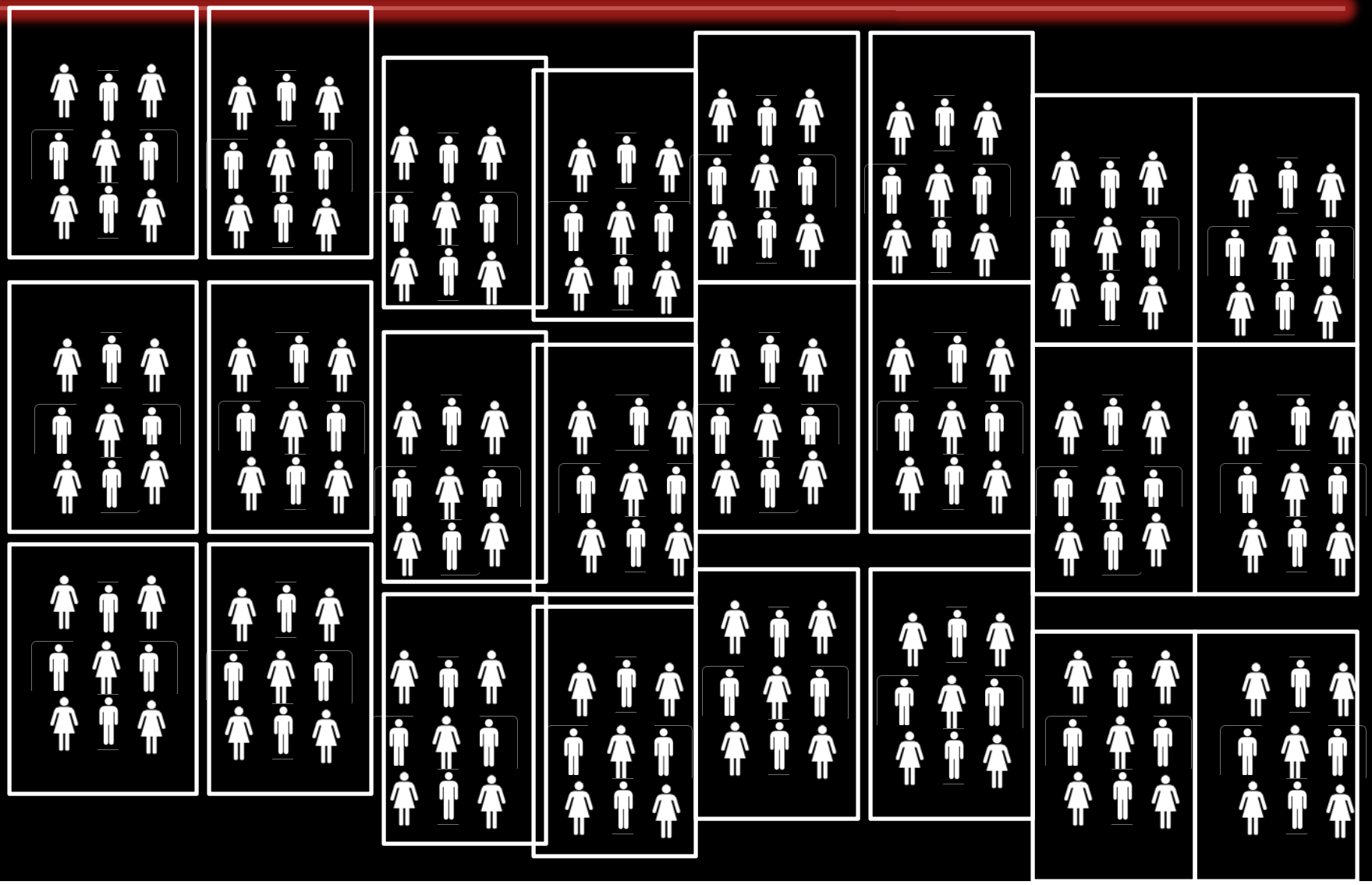
Unidad de aleatorización: 216 individuos



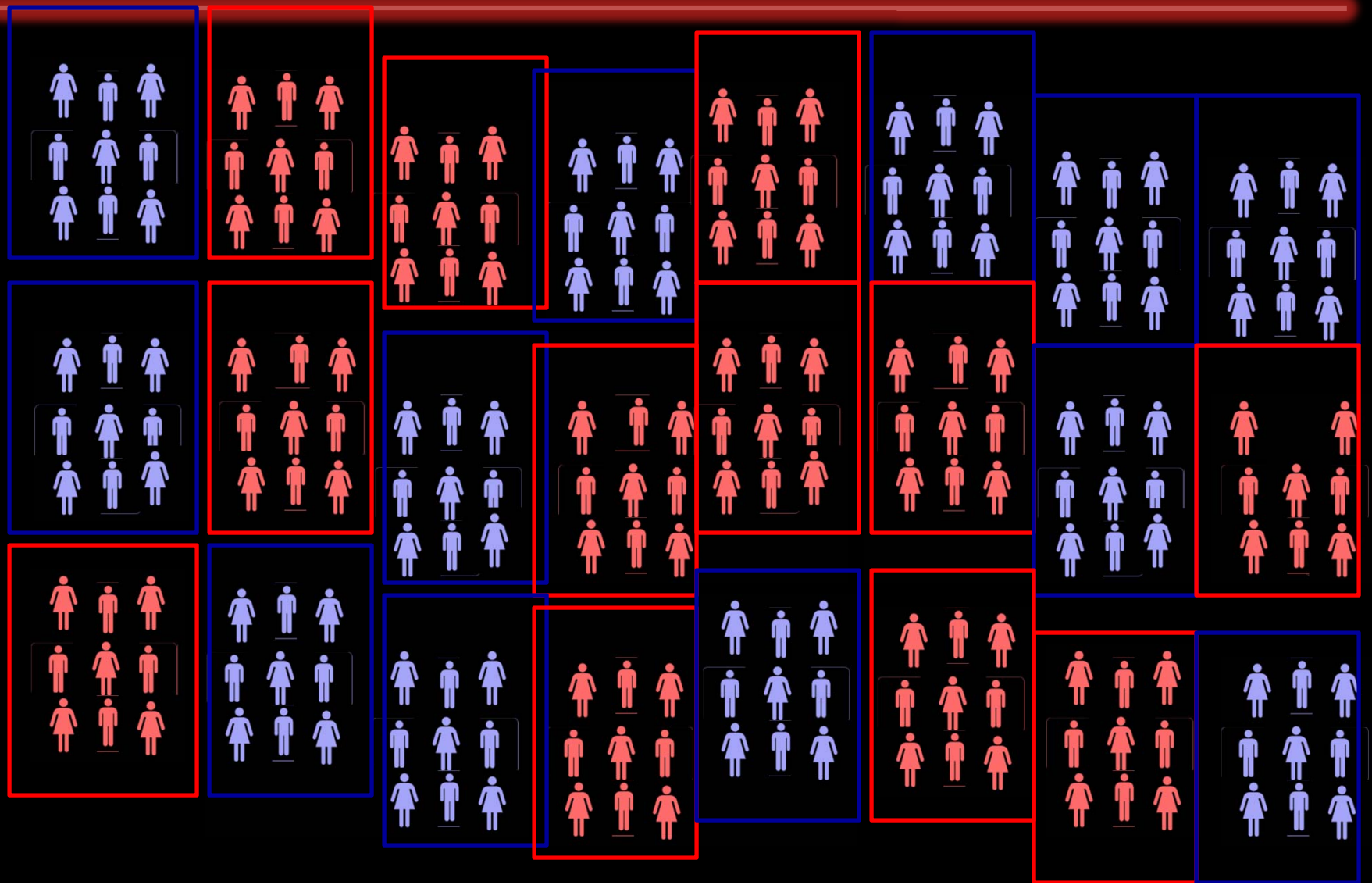
Unidad de aleatorización: 216 individuos



Unidad de aleatorización: unidades agrupadas



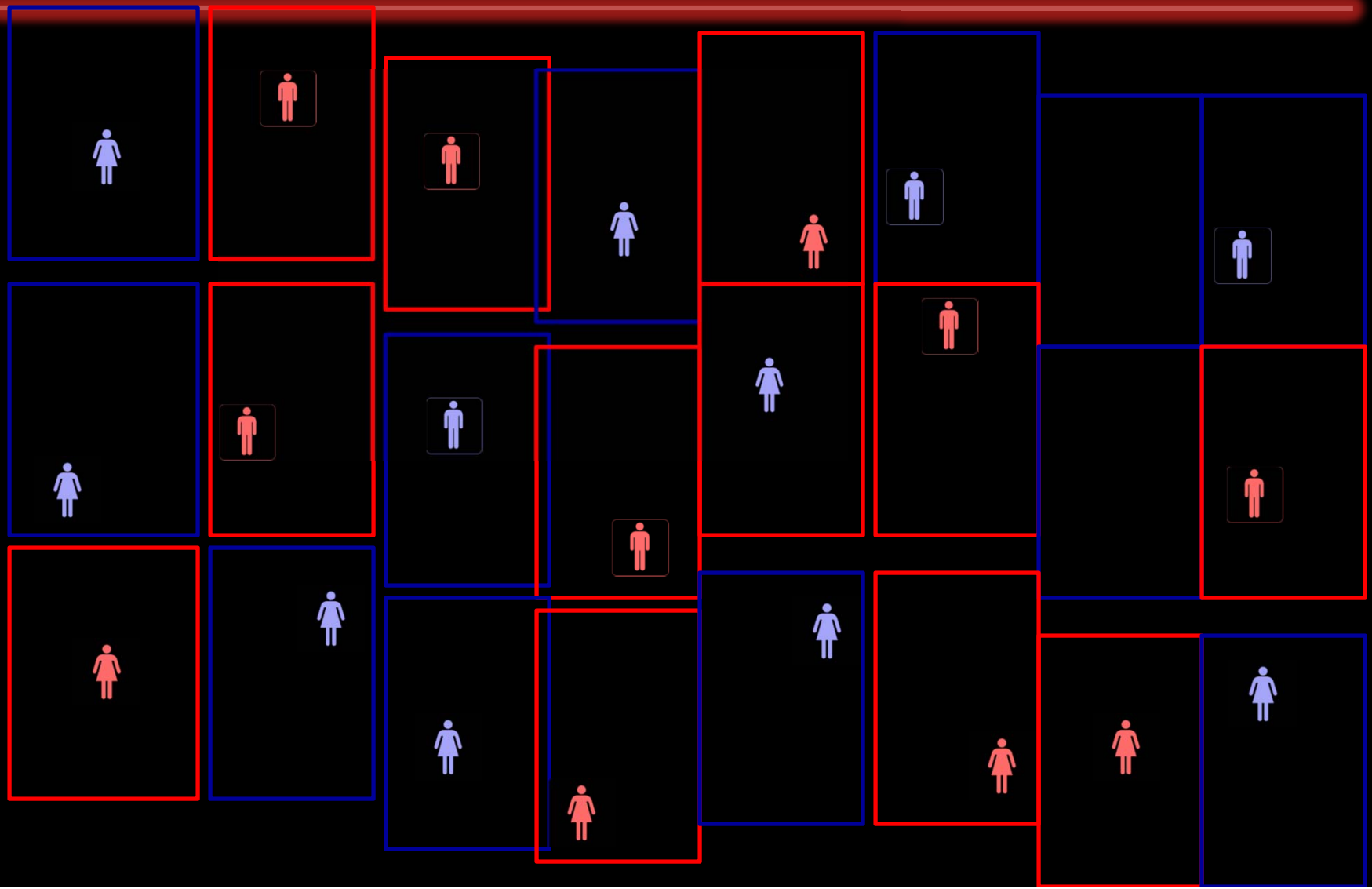
Unidad de aleatorización: 24 unidades agrupadas



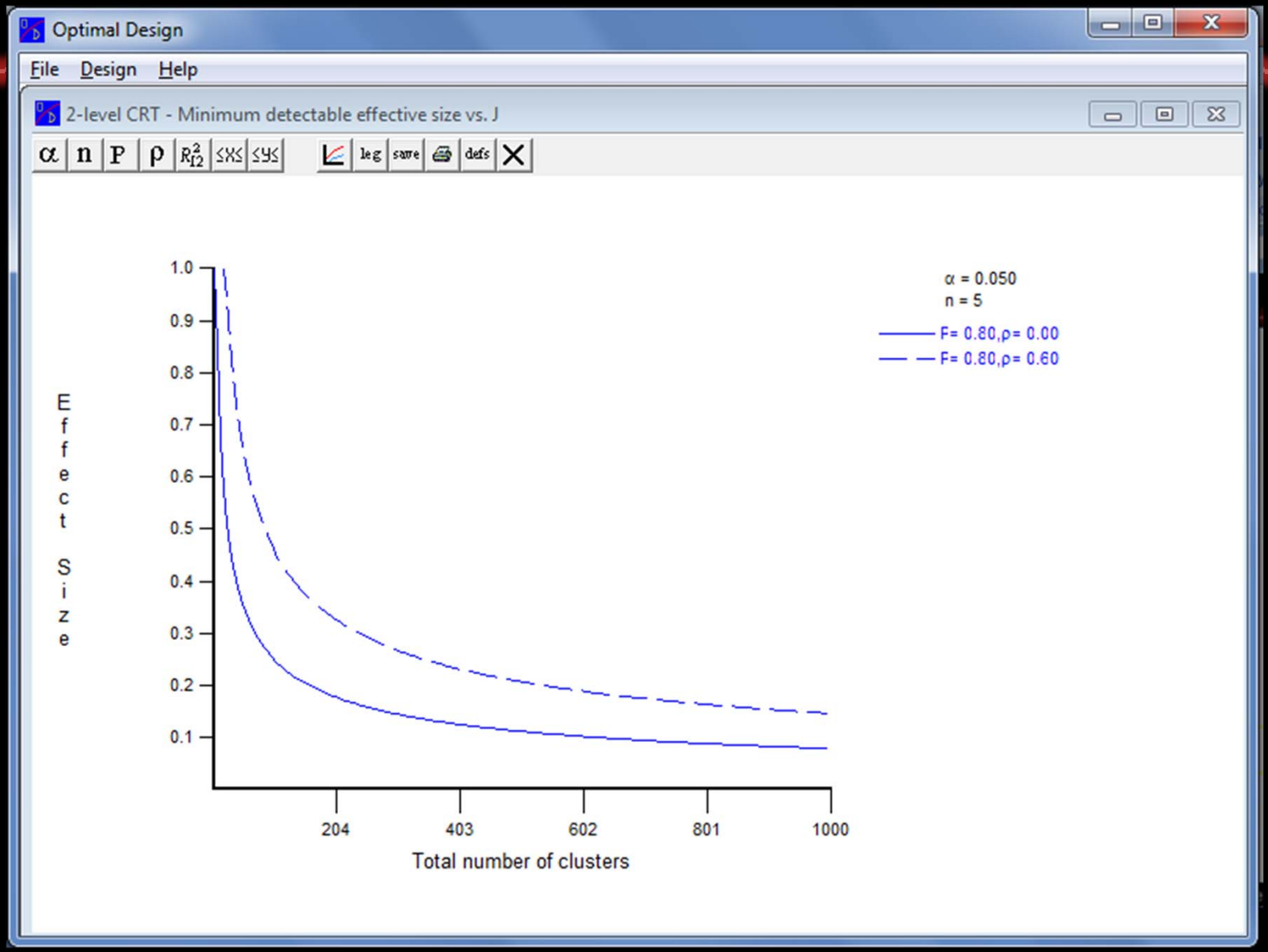
Diseño de conglomerado: intuición

- Muchas razones prácticas para hacer esto:
  - Costos – monetarios y políticos
  - Menor riesgo de contaminación (contagio?) o error administrativo
  - Tipos de intervención (maestros para muchos alumnos, etc.)
- Pero también tiene costos: Si la respuesta se correlaciona dentro de un grupo, se obtiene menos información de la medición de varias personas en el grupo.
- Las estimaciones de *cluster* ajustan por esta correlación – como si fuera un número de observaciones *efectivas*.
- Es más informativo medir a personas no relacionadas
  - En mejor tener 200 encuestas, 2 por conglomerado en 100 conglomerados que 100 por conglomerado en dos conglomerados
- Caso extremo: dentro de cada aglomerado son todos iguales...

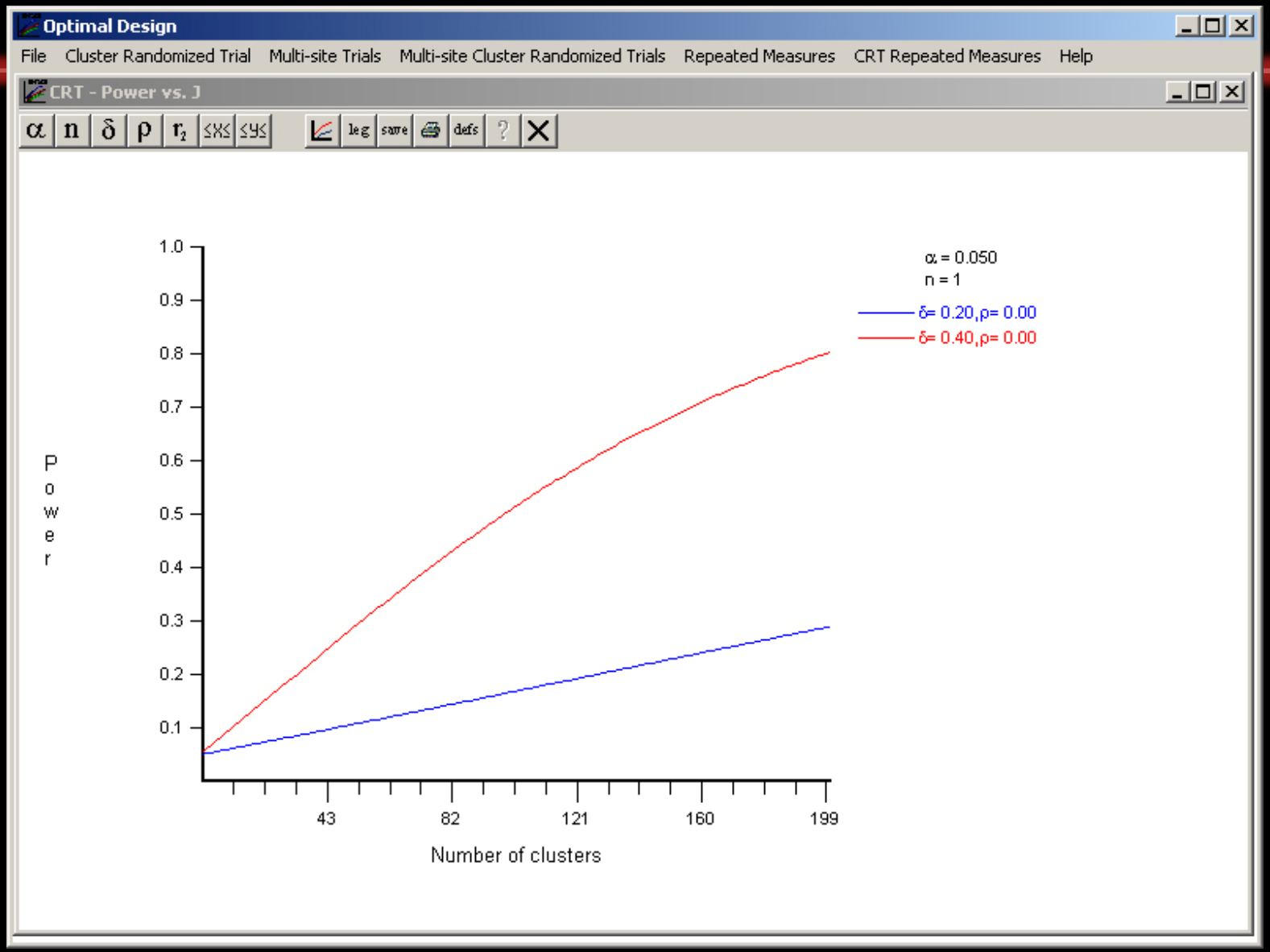
Correlación perfecta intra-clase: 24 unidades…



Poder, efecto detectable y número de grupos



Poder y correlación al interior de los grupos



Valores de r (rho) – correlación intra clase

- Al igual que los porcentajes, r debe estar entre 0 y 1
- Al trabajar con diseños en conglomerados, es más deseable un r menor
- A veces es bajo, 0, 0,05, 0,08, pero puede ser alto: 0,62

Madagascar Matemáticas + Lenguaje	0.5
Busia, Kenia Matemáticas + Lenguaje	0.22
Udaipur, India Matemáticas + Lenguaje	0.23
Mumbai, India Matemáticas + Lenguaje	0.29
Vadodara, India Matemáticas + Lenguaje	0.28
Busia, Kenia Matemáticas	0.62

Algunos ejemplos del tamaño muestral

Estudio	N° de grupos tratamiento/ control	Número total de conglomerados	Tamaño total de la muestra
Empoderamiento de las mujeres	2	Rajasthan: 100 Bengalia Occidental: 161	1996 encuestados 2813 encuestados
Read India de Pratham	4	280 aldeas	17.500 niños
Balsakhi de Pratham	2	Mumbai: 77 escuelas Vadodara: 122 escuelas	10.300 niños 12.300 niños
Programa de Profesores Adicionales en Kenia	8	210 escuelas	10.000 niños
Desparasitación en Kenia	3	75 escuelas	30.000 niños

Consecuencias de los conglomerados

- Los resultados para las personas dentro de un conglomerado pueden estar correlacionados
- Diseño: Debemos tomar en cuenta los conglomerados al planificar el tamaño muestral
- Entre mayor la correlación, se vuelve más importante tener un mayor número de conglomerados en el experimento, dado un número de encuestas fijo.

Consecuencias de los conglomerados

- Es fundamentla establecer un número adecuado de grupos de asignación aleatoria.
- Incluso el número de individuos *al interior* de los grupos a veces importa menos que el número total de grupos.
- Importa el número de *unidades* tratadas.

3. ¿Se cumple total o parcialmente con el protocolo de asignación experimental?

- Tal como los hicimos, cálculos de tamaño muestral asumen que todos los participantes hacen lo que el protocolo de asignación ordena.
- Algunas personas asignadas al tratamiento pueden no tomarlo y algunos asignados a control pueden buscar cómo y recibir tratamiento.
- Por ejemplo, en programa de becas PACES en Colombia, sólo 50% de estudiantes asignados a la beca inicialmente la mantenían 3 años más tarde.

3. ¿Se cumple total o parcialmente con el protocolo de asignación experimental?

- Esto implica que mínimo tamaño muestral para detectar un impacto dado debe ser CUATRO veces mayor
- En general tamaño muestral mínimo incrementa con el recíproco del cuadrado de la diferencia en cumplimiento entre tratamiento y control: $1/(c-s)^2$

ALGUNOS ASPECTOS PRÁCTICOS

Errores comunes y tentaciones

- La tragedia de un buen programa con buena estrategia de evaluación con problemas de poder:
  - Tamaño de muestra
  - Sobre-optimismo en magnitud de efectos
  - Pocos aglomerados – mala estimación de rho
- Cada vez menos común – instituciones donantes y de evaluación requieren cálculos de poder.
- Tentaciones: ¿Y si probamos 25 tratamientos diferentes?

Leer el manual de instrucciones

- Por suerte tenemos el “toolkit” (Duflo et al.) y mucha experiencia en ciencias sociales, epidemiología – no hay excusas para los errores más sencillos.
- Pero muchas veces nuestros programas tendrán efectos esperados pequeños, y/o tendremos fondos limitados.
- Aprovechar la interacción de los factores que vimos hoy (clusters, covariables, etc.) para maximizar el poder de las evaluaciones mediante diseños muestrales más sofisticados. PERO NO HAY MAGIA.
- Poder es planificar!

Riesgo para la salud de su evaluación

Ante cualquier duda consulte a su estadístico/a amigo/a

Objetivos de la clase

1. Refrescar/entender elementos básicos de estadística:
   - Estimación, Prueba de hipótesis, Significancia estadística
2. Concepto clave en evaluaciones aleatorias: Poder y los factores que lo influyen
   - Tamaño muestral, efectos mínimos detectables
   - Aglomerados y otros factores
3. Asignar recursos escasos en una evaluación.
   - Maximizar la probabilidad de encontrar efectos de programas que funcionan.

Gracias!

Guillermo Cruces

povertyactionlab.org